A course of mathematics

Book II

Construction of complex numbers and functions. Basic mathematical analysis.

Naumov Ivan

Book II. Construction of complex numbers and functions.
Basic mathematical analysis.

ISBN 978-952-94-0574-9
ISBN 978-952-94-0575-6
ISBN 978-952-94-0576-3
ISBN 978-952-94-0577-0

Illustrator, designer
Gudalova Tatiana

Finland 2018

Preface

This book continues the course of mathematics. We will start from the construction of vectors and complex numbers. And then we will continue with a course of basic mathematical analysis, which will be used as a tool for construction of power functions and radians. Both power functions and radians are usually defined very briefly, without any detailed explanations. But some explanations actually must be done. Consider the simple function $f(x)=2^x$. What is the value of this function at the point $x=\sqrt[3]{3}$? How can we find the number $2^{\sqrt[3]{3}}$? It's not a problem to understand how to raise real numbers to integer or even to rational powers, but it's not enough. If we want to consider the function $f=2^x$ as a function of a **real** variable x, we need to understand precisely what is a value of this function at any irrational number. Put simply, we need to define the numbers like $2^{\sqrt[3]{3}}, 2^{\pi},\ (\sqrt[2]{5})^{\sqrt{8}},\ \sqrt[4]{5}^{\sqrt[7]{8}}$. The notion of radians is also extremely important, I've never come across any books that provide some rigorous explanation of it. There is always just a brief definition, where radians are defined through the notion of a length of an arc of a unit circle, but there is no explanation what is a length of such arc. How do we actually measure a length of an arc? Does there exist an arc of a circle with a length $\sqrt{5}$? Usually people like to say we can straighten an arc and then use a ruler. At first, it's not mathematics, and secondly, there is no ruler that can measure lengths like $\sqrt{5},\ 2/e,\ \sqrt[3]{3}$. And as a result there is no normal definition of a length of an arc, which leads to the fact that the definition of a radian argument is based on nothing, so it is not a definition. That's why I decided to provide a complete explanation of this notion.
As far as I'm concerned, it's always very important to specify the most important notions, because if we don't have any "foundation" (like definitions and rules to follow), it's easy to make a wrong reasoning and to finish with a wrong result. Really, many things that seem obvious to us in reality appear not only not obvious, but even wrong. I'll give the simple example. Suppose two guys compete in arm wrestling. Most of people will agree that one guy is winning because his palm exerts a greater force on the palm of the other guy. So his palm pushes the other palm stronger than the other palm pushes his palm, and that's why he is winning. It sounds like an obvious explanation, but it is completely wrong. According to the basic laws of nature (3-rd Newton's law), at any given moment, the forces, which palms exert on each other, are precisely equal in magnitude. So at any given moment, palms push each other with equal forces. And something that seemed "obvious" turned out to be wrong. And there are a lot of examples like this one. That's why it's always very important not to rely solely on our intuition, but stick to the basic laws and definitions, which are precisely defined.

It is initially planned to provide a complete course of mathematics, which consists of 8 books. 1-st, 2-nd and 3-rd book are dedicated to explain the foundations of mathematics, and the final 6,7,8 books will include more advanced subjects, like mechanics, differential equations and equations of mathematical physics.

All The Best, Naumov Ivan.

naumov8v@gmail.com

Table of contents.

I express my sincere gratitude to:

Galochkin Alexandr Ivanovich

Drutsa Alexey Valerievich

Kozko Artem Ivanovich

Because their lectures and seminars made me interested in math
and formed my approach to this subject.

I also would like to thank **Alexander Bebris**, as his educational web site and the project "English Galaxy" provided an amazing course of English grammar, which is essential for writing.

10

Vectors

Vectors

In this chapter we will "build" vectors by using the basic principles and axioms of stereometry as a foundation. It's much more common to use more "intuitive" approach to vectors, where most of the properties of vectors are taken for granted, or "proved" by a picture, which usually reflects only a particular case. Such approach also uses the term "vectors have the same direction", without defining what is a same direction of parallel vectors. And as a result, everything depends on the eye of the observer, which is not a good thing. We will stick to another approach, that doesn't have such flaws, but in return we will have to "pay" the price for it, e.g., in some cases we will have to pay more attention to details and consider several possible variants. Let's start.

Def1. Let AB is any segment in the space. Let's choose a "direction" of AB. For example, let A is a "start point" and B is an "end point". The end point B must be marked like an arrowhead **[pict1]**. Now we have the object which must be denoted as $\overrightarrow{AB}$. The small arrow above the letters shows us that A is a start point and B is an end point.
$\overrightarrow{AB}$ (an arrow in the space) is called a "vector".

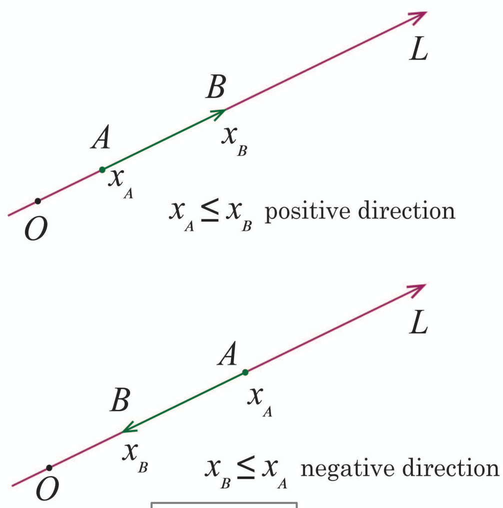
pict.1

The length of the segment AB is called a module of $\overrightarrow{AB}$, and we write $\left|\overrightarrow{AB}\right| = AB$.

If points A, B coincide, so $A = B$, then we say that $\overrightarrow{AB}$ is a zero vector, and we write $\overrightarrow{AB} = \vec{0}$.

Next, let's fix any line L in the space and turn it into a coordinate line.
If we move along L in a certain direction, we always meet greater numbers, than we met before.
Such direction is called "a positive direction" of an axis.
Any axis in the space must be depicted as a segment with a certain direction on it, which indicates a positive direction of an axis. Let $\overrightarrow{AB}$ lies on some axis L (it means that $A, B \in L$).

We say that $\overrightarrow{AB}$ has a positive direction if $x_A \le x_B$ (the coordinate of A is not greater than the coordinate of B) **[pict2]**, and we say that $\overrightarrow{AB}$ has a negative direction if $x_B \le x_A$. Any zero vector $\vec{0}$, which lies on L, has a positive and a negative direction at the same time.



pict.2

Two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are called collinear if they both lie on the same line L.
Vectors, that do not lie on the same line, are called not collinear.

Def2. Let $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are **collinear** vectors, if they both have the same direction (both positive, or both negative) then we say that these vectors are **co**-directed. If collinear vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ have different directions, then we say that these vectors are **anti**-directed.

We will say that two lines are parallel if **[A]** or **[B]**:

[A] there exist some plane which contains both lines, and these lines do not have any common points.

[B] lines coincide: these lines are in fact the same line.

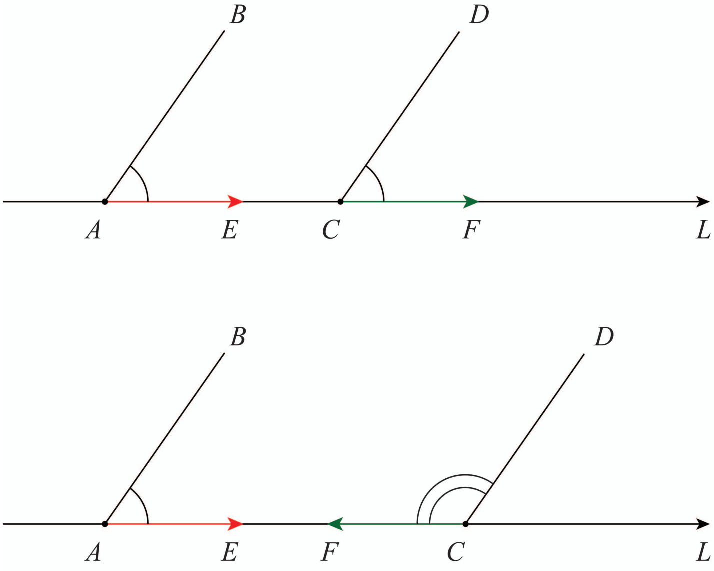
Exercise1. Lines AB and CD lie on the same plane Π, and they intersect some line $L \in \Pi$ at points A and C. Vectors $\overrightarrow{AE}$ and $\overrightarrow{CF}$ are collinear (and lie on L) **[pict3.1]**.

If vectors $\overrightarrow{AE}$ and $\overrightarrow{CF}$ are co-directed, then

[A] $\angle BAE = \angle DCF \Leftrightarrow AB \parallel CD$

If vectors $\overrightarrow{AE}$ and $\overrightarrow{CF}$ are anti-directed, then

[B] $\angle BAE + \angle DCF = 180^o \Leftrightarrow AB \parallel CD$

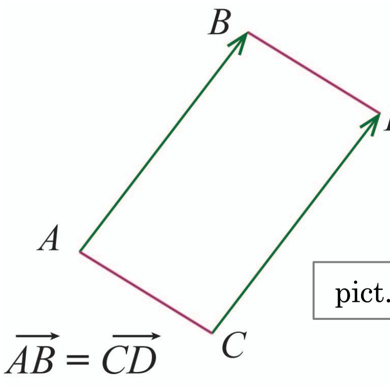


pict.3.1

Put simply. Lines are on one plane are parallel if and only if they form equal angles with co-directed collinear vectors. And lines on the plane are parallel if and only if they form angles, which sum is 180^o, with anti-directed collinear vectors.

Def3. Two **not collinear** vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal **if** $AB = CD$ and $AB \parallel CD$ and $AC \parallel BD$ (let's call these requirements **[S]**)**[pict3]**. **Note**: from **[S]** immediately follows that $ABCD$ is a parallelogram and also $AC = BD$. Two **collinear** vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal if $AB = CD$ and these vectors are co-directed.

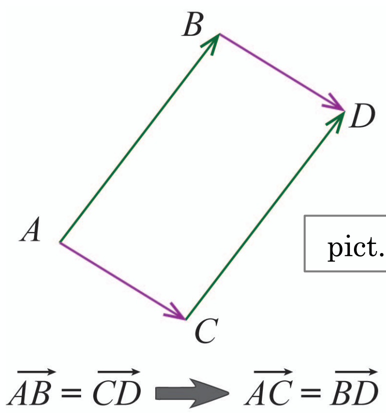
Note. When we write $BD \parallel AC$ we mean that the line, which passes through B and D, is parallel to the line which passes through A and C. We can also say "vectors are parallel" if vectors lie on parallel lines, and we can write $\overrightarrow{AB} \parallel \overrightarrow{CD}$.



pict.3

From the definition of vector equality **[S]**: two not collinear vectors are equal if their modules are equal, they are parallel, and lines, which connect their start and end points, are parallel.

Assertion1. Two not collinear vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal $\Leftrightarrow$ The vector $\overrightarrow{AC}$ which connects their start points is equal to the vector $\overrightarrow{BD}$ which connects their end points.



pict.4

Proof. $\Rightarrow$ Let $\overrightarrow{AB} = \overrightarrow{CD}$, by definition it means that $AB = CD$ and $AB \parallel CD$ and $AC \parallel BD$, from here follows that $ABCD$ is a parallelogram and $AC = BD$.

We want to prove that $\overrightarrow{AC} = \overrightarrow{BD}$, so we need to show **[S]**,

which is obviously true, really $AC = BD,\ \ AC \| BD,\ \ AB \| CD$. Therefore $\overrightarrow{AC} = \overrightarrow{BD}$. **[pict4]**.

$\Leftarrow$Let now $\overrightarrow{AC} = \overrightarrow{BD}$, from here, in the exactly similar way as above, we can get $\overrightarrow{AB} = \overrightarrow{CD}$.

Auxiliary1. Collinear vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal $\Leftrightarrow x_B - x_A = x_D - x_C$.

$\Rightarrow$Let collinear vectors are equal $\overrightarrow{AB} = \overrightarrow{CD}$. **[A]** If both these vectors have a positive direction, then $x_A \le x_B$ and $x_C \le x_D$. The length of the segment AB is equal to $x_B - x_A$ and the length of the segment CD is equal to $x_D - x_C$. Vectors are equal, then $AB = CD \Rightarrow x_B - x_A = x_D - x_C$.

[B] If both vectors have negative direction, then $x_A \ge x_B$ and $x_C \ge x_D$. The length of AB is $x_A - x_B$, and the length of CD is $x_C - x_D$. As vectors are equal, then $AB = CD \Rightarrow$
$x_A - x_B = x_C - x_D \Leftrightarrow x_B - x_A = x_D - x_C$.

Conversely. $\Leftarrow$ $x_B - x_A = x_D - x_C$. **[A]** If $x_A \le x_B$, then there must be $x_C \le x_D$. Then both vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ have a positive direction (so these vectors are co-directed), and $x_B - x_A$ is a length of AB and $x_D - x_C$ is a length of CD, then $AB = CD$. Therefore $\overrightarrow{AB} = \overrightarrow{CD}$.

[B] If $x_A \ge x_B$ then there must be $x_C \ge x_D$. Then both $\overrightarrow{AB}$ and $\overrightarrow{CD}$ have a negative direction. And $x_A - x_B$ is a length of AB and $x_C - x_D$ is a length CD. From $x_B - x_A = x_D - x_C$ follows that $x_A - x_B = x_C - x_D \Rightarrow AB = CD$. Then $\overrightarrow{AB} = \overrightarrow{CD}$.

We can use auxiliary1 to prove the assertion2 which is very similar to the assertion1.

Assertion2. Two collinear vectors are equal $\Leftrightarrow$ vector that connects their start points is equal to the vector that connects their end points

Proof. $\Rightarrow$Let $\overrightarrow{AB} = \overrightarrow{CD}$ are equal collinear vectors, then $x_B - x_A = x_D - x_C$ **[A]**.

We have to show that vectors $\overrightarrow{AC}$ and $\overrightarrow{BD}$ are equal, by auxiliary1 they are equal if and only if $x_C - x_A = x_D - x_B$- this equality immediately follows from **[A]**.

$\Leftarrow$vectors, which connects start and end points of $\overrightarrow{AB}$ and $\overrightarrow{CD}$, are equal, it means that $\overrightarrow{AC} = \overrightarrow{BD}$. Then $x_C - x_A = x_D - x_B$- from this equality follows the equality **[A]** and therefore $\overrightarrow{AB} = \overrightarrow{CD}$.

We can generalize, from the assertions 1,2 follows the

Theorem1. Two vectors are equal $\Leftrightarrow$ vector that connects their start points is equal to the vector that connects their end points

Theorem2. Vector equality "=" is symmetrical: $\overrightarrow{AB} = \overrightarrow{CD} \Leftrightarrow \overrightarrow{CD} = \overrightarrow{AB}$.

Proof: For not collinear equal vectors: $\overrightarrow{AB} = \overrightarrow{CD}$, then $AB = CD$ and $AB \| CD$ and $AC \| BD$ and also $AC = BD$. If we want to show that $\overrightarrow{CD} = \overrightarrow{AB}$ (by **[S]**) there must be:

$CD = AB,\ CD \| AB,\ CA \| DB$- and we already have it. Let now $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are collinear and equal, then there is nothing to prove, symmetricity follows right from the definition, we can also write $\overrightarrow{AB} = \overrightarrow{CD} \Rightarrow x_B - x_A = x_D - x_C \Rightarrow x_D - x_C = x_B - x_A \Rightarrow \overrightarrow{CD} = \overrightarrow{AB}$. We showed that $\overrightarrow{AB} = \overrightarrow{CD} \Rightarrow \overrightarrow{CD} = \overrightarrow{AB}$. The proof of the converse assertion is exactly similar.

Theorem3. Vector equality "=" is transitive: $\overrightarrow{AB} = \overrightarrow{CD}$ and $\overrightarrow{CD} = \overrightarrow{EF}$, then $\overrightarrow{AB} = \overrightarrow{EF}$.

Comment. Despite the fact that this theorem looks very simple, it appears the most "demanding" theorem in the vector construction. It happens because we have precisely defined what are equal vectors, that's why we have very strict and clear requirements to vectors in order them to be equal. Remember that if we don't stick to the initial definitions, all the process loses it's value.

Simple exercise. For any vectors $\overrightarrow{AB} = \overrightarrow{CD} \Leftrightarrow \overrightarrow{BA} = \overrightarrow{DC}$.

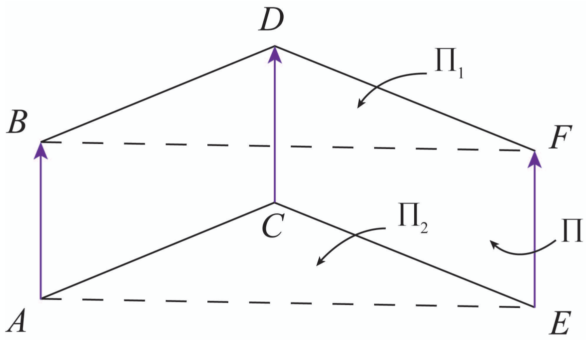
Comment. If one of vectors $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$ is a zero vector $\vec{0}$, then the proof is obvious (each vector must be a zero vector in such case). From now on, there are no zero vectors among $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$. Also, if some two of vectors $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$ coincide, i.e., some two of these vectors are in fact the same vector, then there is nothing to prove. From now on, there are no coinciding vectors among $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$.

Proof. Let there are no collinear vectors among $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$ **[pict5]**. We have to show that $\overrightarrow{AB} = \overrightarrow{EF}$, by definition, it means that we need to prove that:
$AB = EF$ and $AB \| EF$ and $AE \| BF$.

From $\overrightarrow{AB} = \overrightarrow{CD}$ follows that $AB = CD,\ AB \| CD$.

From $\overrightarrow{CD} = \overrightarrow{EF}$ follows that $CD = EF,\ CD \| EF$.

Therefore $AB = EF$ and $AB \| EF$.



pict.5

The last thing that we need to show: $AE \| BF$. As $AB \| EF$, both these segments lie on the same plane Π, this plane also contains the segments AE, BF **[T]**. We have to show that these two segments are not diagonals of the parallelogram $ABEF$ (we can't use **[pict5]** as an argument, there must be a normal proof), we will show that the segments **[T]** lie in two parallel planes Π_1 and Π_2 and therefore they don't have any common points (then **[T]** must be two opposite sides of the parallelogram $ABEF$, then **[T]** are parallel and everything is proved). From $\overrightarrow{AB} = \overrightarrow{CD}$ follows that $BD \| AC$ and from $\overrightarrow{CD} = \overrightarrow{EF}$ follows that $DF \| CE$. Let's draw the plane Π_1 through the lines BD and DF and the plane Π_2 through the lines AC and CE. These planes must be parallel $\Pi_1 \| \Pi_2$(because $BD \| AC$ and $DF \| CE$). The plane Π_1 contains B, D, F, then $BF \in \Pi_1$, the plane Π_2 contains A, C, E, then $AE \in \Pi_2$. Then **[T]** lie in two parallel planes.

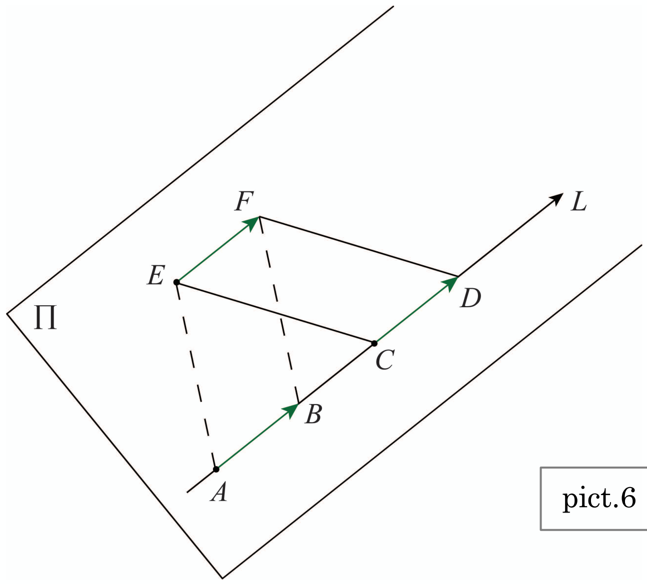
If all three vectors $\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{EF}$ are collinear, then the proof is obvious:

$\overrightarrow{AB}=\overrightarrow{CD}\Rightarrow x_B-x_A=x_D-x_C \,||\, \overrightarrow{CD}=\overrightarrow{EF}\Rightarrow x_D-x_C=x_F-x_E$.

Then there must be: $x_B-x_A=x_F-x_E\Rightarrow\overrightarrow{AB}=\overrightarrow{EF}$.

Let now some two of three vectors $\overrightarrow{AB},\overrightarrow{CD},\overrightarrow{EF}$ are collinear. There are exactly 3 different variants (if 2 of these 3 vectors are collinear).

Let $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are collinear **[pict6]** $\Rightarrow$ they lie on the same line L. So we have $\overrightarrow{AB}=\overrightarrow{CD}$ and also $\overrightarrow{CD}=\overrightarrow{EF}$. We want to show that $\overrightarrow{AB}=\overrightarrow{EF}$.



pict.6

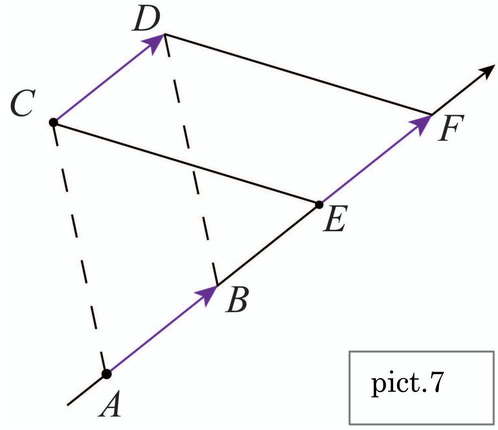
As $\overrightarrow{CD}=\overrightarrow{EF}$, then $\overrightarrow{CD}\,||\,\overrightarrow{EF}\Rightarrow L\,||\,\overrightarrow{EF}$ and we can draw the plane Π through the line L and the vector $\overrightarrow{EF}$, then Π contains all three vectors $\overrightarrow{AB},\overrightarrow{CD},\overrightarrow{EF}$.

In the exactly similar way as in the previous reasoning we deduce: $AB=EF$ and $AB\,||\,EF$. The last thing that we need to show: $AE\,||\,BF$ (then $\overrightarrow{AB}=\overrightarrow{EF}$ by definition). From $\overrightarrow{CD}=\overrightarrow{EF}$ follows that $CE\,||\,DF$ and $CE=DF$. From $\overrightarrow{AB}=\overrightarrow{CD}$ follows that vector which connects their start points is equal to the vector which connects their end points: $\overrightarrow{AC}=\overrightarrow{BD}$ (collinear vectors) then also (simple exercise from above) $\overrightarrow{CA}=\overrightarrow{DB}$.

Let's sum up: $CE=DF$ and $CE\,||\,DF$ and $\overrightarrow{CA}=\overrightarrow{DB}$. Lines CE and DF are parallel, then (exercise1) they form equal angles with collinear co-directed vectors $\overrightarrow{CA}$ and $\overrightarrow{DB}$, so $\angle ACE=\angle BDF$. Then we have: $CE=DF$ and $AC=BD$ (because $\overrightarrow{CA}=\overrightarrow{DB}$) and $\angle ACE=\angle BDF$, therefore $\Delta ACE=\Delta BDF$. From here follows $\angle EAC=\angle FBD$ (as these angles lie in front of equal sides in equal triangles). Also $\overrightarrow{AC}$ and $\overrightarrow{BD}$ have the same direction (because they are equal), then (exercise1) there must be $EA\,||\,FB$ and everything is proved.

The case when $\overrightarrow{CD}$ and $\overrightarrow{EF}$ are collinear is exactly similar to the case that is considered above.

The last case that we need to consider: $\overrightarrow{AB}$ and $\overrightarrow{EF}$ are collinear (they lie on the same line) and $\overrightarrow{CD}$ doesn't lie on that line **[pict7]**. From $\overrightarrow{AB}=\overrightarrow{CD}$ follows that $AC\,||\,BD,\ AC=BD$.



pict.7

From $\overrightarrow{CD}=\overrightarrow{EF}$ follows that $CE\,||\,DF,\ CE=DF$.

Let's show that $\Delta ACE=\Delta BDF$:

we have $AC=BD,\ CE=DF$ **[G]** (we also need equal angles).

Comment: we can't just say that $\angle ACE$ and $\angle BDF$ are equal because the sides of one angle are parallel to the sides of the other angle. In such case angles mustn't be equal. If sides of one angle are parallel to the sides of the other angle, these angles either equal, or their sum is 180^o (and they mustn't be equal in such case). So we need to provide a normal reasoning, based on our assumptions, which we made at the very beginning of this chapter. Let's consider two possible cases **[A]** and **[B]**.

[A] Let $\overrightarrow{AE}$ and $\overrightarrow{BF}$ are co-directed. As $AC \parallel BD$, then (exercise1) these lines form equal angles with collinear co-directed vectors $\overrightarrow{AE}$ and $\overrightarrow{BF}$, and we have: $\angle CAE = \angle DBF$.

As $\overrightarrow{AE}$ and $\overrightarrow{BF}$ are co-directed, vectors $\overrightarrow{EA}$ and $\overrightarrow{FB}$ are also co-directed, and we have parallel lines $CE \parallel DF$, then (exercise1) $\angle CEA = \angle DFB$.

Let's sum up: for ΔACE and ΔBDF we have **[G]** and we have just deduced that the angles near the base AE (in ΔACE) are equal to the angles near the base BF (in ΔBDF), as the sum of angles of any triangle is 180^o, there must be $\angle C = \angle D$ (or the same $\angle ACE = \angle BDF$), therefore $\Delta ACE = \Delta BDF$. As these triangles are equal, their bases are also equal: $AE = BF$, we know that $\overrightarrow{AE}$ and $\overrightarrow{BF}$ are co-directed, therefore there must be $\overrightarrow{AE} = \overrightarrow{BF}$ (by definition), then (theorem1) vectors which connect their start and end points are also equal, i.e., $\overrightarrow{AB} = \overrightarrow{EF}$

it is exactly what we need.

[B] Let $\overrightarrow{AE}$ and $\overrightarrow{BF}$ are anti-directed. We have $AC \parallel BD$ then (exercise1) there must be:

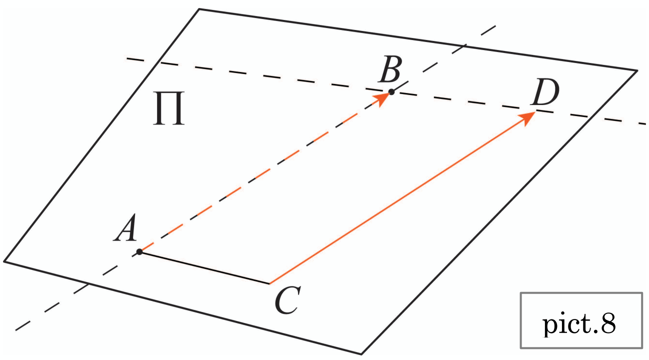
$$\angle CAE + \angle DBF = 180^o \Leftrightarrow \angle CAE = \alpha, \ \angle DBF = 180^o - \alpha.$$

As vectors $\overrightarrow{AE}$ and $\overrightarrow{BF}$ are anti-directed, vectors $\overrightarrow{EA}$ and $\overrightarrow{FB}$ must be also anti-directed, and we have $CE \parallel DF$, then (exercise1) $\angle CEA + \angle DFB = 180^o \Leftrightarrow \angle CEA = \beta, \ \angle DFB = 180^o - \beta$.

Then in ΔACE and ΔBDF the sum of all 4 angles which lie near the bases AE, BF is 360^o, it is impossible because the sum of all 6 angles of these triangles is 360^o. In this case there must be $\angle ACE = 0^o, \ \angle BDF = 0^o \Rightarrow A = E, \ B = F$, then $\overrightarrow{AB}$ and $\overrightarrow{EF}$ coincide, so they are co-directed, which contradicts to **[B]**. Then the case **[B]** is impossible. The theorem is proved.

Theorem4. For any point A in the space and any vector $\overrightarrow{CD}$ there exist only one vector $\overrightarrow{AB}$ (which start point is A) that is equal to $\overrightarrow{CD}$.

Proof. Existence. If $\overrightarrow{CD}$ is a zero vector, then $C = D$, then we take the vector $\overrightarrow{AA}$ which is also a zero vector. Let $\overrightarrow{CD}$ is not a zero vector. If A and $\overrightarrow{CD}$ do not lie on the same line L, then we can draw the plane Π through A, C, D [pict8]. And (in Π) we can build



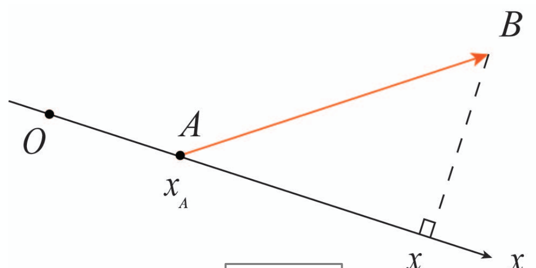
pict.8

a parallelogram $ABCD$: $AB \parallel CD$, $AC \parallel BD$. Then $\overrightarrow{AB}$ is (obviously) the vector we need. If A and $\overrightarrow{CD}$ belong to some line L, then we turn L into the axis. Let's choose the positive direction on L in such way that $\overrightarrow{CD}$ has a positive direction. Then we can take the point $B \in L$ such that $AB = CD$ and $\overrightarrow{AB}$ has a positive direction. Then $\overrightarrow{AB} = \overrightarrow{CD}$.

Uniqueness. Let's assume that $\overrightarrow{AB} = \overrightarrow{CD}$ and $\overrightarrow{AB_1} = \overrightarrow{CD}$. By transitivity (**theorem3**) we have $\overrightarrow{AB_1} = \overrightarrow{AB}$ **[V]**. From $\overrightarrow{AB_1} = \overrightarrow{AB}$ follows (**theorem1**) that $\overrightarrow{AA_1} = \overrightarrow{BB_1} \Leftrightarrow \vec{0} = \overrightarrow{BB_1} \Rightarrow B = B_1$ and vectors $\overrightarrow{AB_1}$, $\overrightarrow{AB}$ coincide.

Coordinates. Let's fix some plane Π and an axis Ox on this plane. We will consider only vectors $\overrightarrow{AB} \in \Pi$ which start point A belongs to Ox.

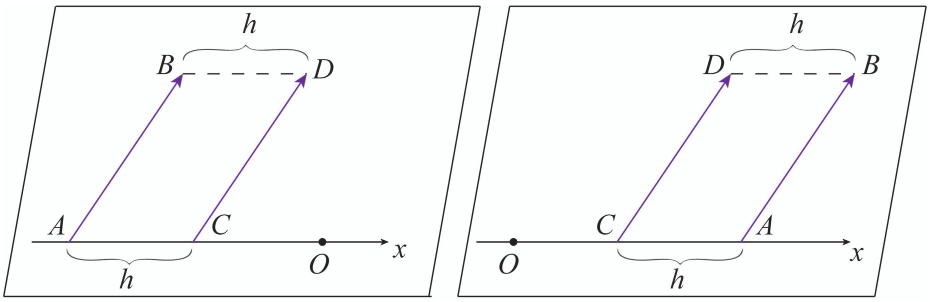
For any vector $\overrightarrow{AB}$, the difference $x_B - x_A$ (where x_B is a coordinate of B and x_A is a coordinate of A) is called an x-coordinate of $\overrightarrow{AB}$ **[pict9]**.
So, for any vector, the difference of it's end point coordinate and it's start point coordinate is an x-coordinate of a vector.



pict.9

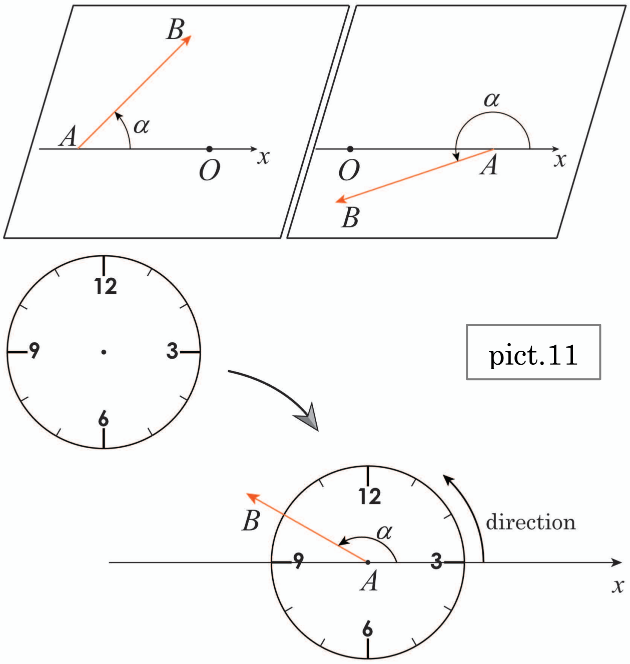
Property1. Equal vectors have equal x-coordinates: if $\overrightarrow{AB} = \overrightarrow{CD}$, then $x_B - x_A = x_D - x_C$.

Proof. From $\overrightarrow{AB} = \overrightarrow{CD}$ follows that $ABDC$ is a parallelogram, where $BD \parallel AC \Rightarrow BD \parallel Ox$ (do not forget that we consider only vectors which start point belongs to Ox). Let h is a length of the side BD. Then $x_D = x_B + h$ and $x_C = x_A + h$ **[pict10]**, or vice versa $x_B = x_D + h$ and $x_A = x_C + h$. In any case we can subtract our equalities and get: $x_B - x_A = x_D - x_C$.



pict.10

Def4. The angle of inclination of any vector $\overrightarrow{AB}$ is the angle $\alpha \in [0^o, 360^o)$ between $\overrightarrow{AB}$ and Ox, which is counted from the positive direction of Ox in the counterclockwise direction **[pict11]**.



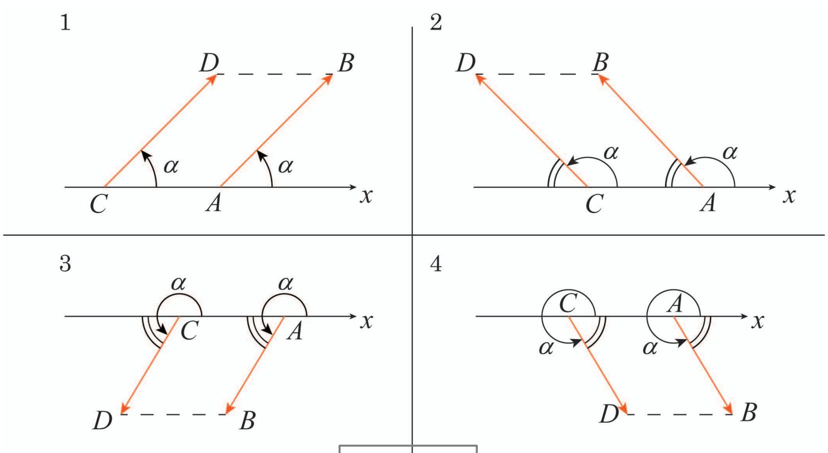
pict.11

Let's pay more attention to the **def4**, because it has a great importance in math.
We need to have some standard according to which we can always understand how to determine the angle between a vector and an axis.
And our "standard" is an ordinary clock dial **[pict11]**.

Let we have some vector $\overrightarrow{AB}$ and some axis Ox, where $A \in Ox$. The dial-center must be placed at A such that the direction from 9 to 3 is a positive direction of Ox, then the direction of rotation from 3 to 9 is our direction, and we must calculate the angle $\alpha \in [0^o, 360^o)$ (between $\overrightarrow{AB}$ and Ox), by moving from Ox in this direction.

Property2. Equal vectors have equal angles of inclination.

Proof. Let $\overrightarrow{AB} = \overrightarrow{CD}$, if $A = C$ then vectors $\overrightarrow{AB}, \overrightarrow{CD}$ coincide and there is nothing to prove. Let $A \neq C$, then $ABDC$ is a parallelogram and $BD \| AC \Rightarrow BD \| Ox$.
Without loss of generality, let $x_A > x_C$.
Let's just consider separately all the possible cases for the angle of inclination α of $\overrightarrow{AB}$. So: $\alpha \in (0^o, 90^o)$, $\alpha \in (90^o, 180^o)$ $\alpha \in (180^o, 270^o)$, $\alpha \in (270^o, 360^o)$ **[pict12]**. In each case parallel lines AB and CD form equal **acute** angles with Ox. By using these equal acute angles in each case, we can easily check that the angles of inclination of $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are really equal.



pict.12

And finally, the cases $\alpha = 0^o$, $\alpha = 90^o$, $\alpha = 180^o$, $\alpha = 270^o$ are obvious.

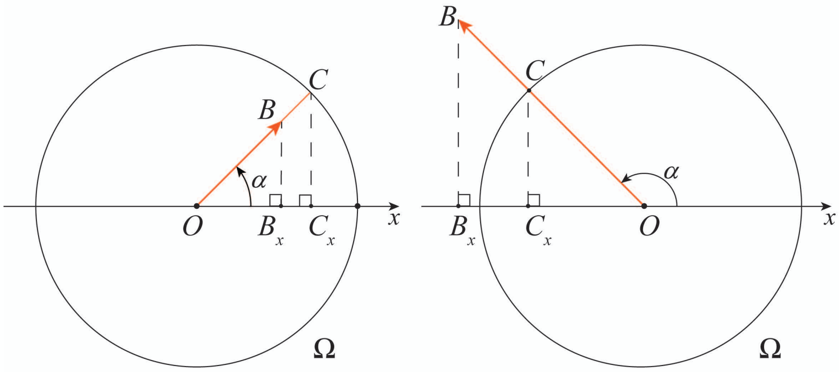
Def. Any vector $\overrightarrow{OB}$, which start point is located at the origin O, is called a radius vector (the origin O is a point with a zero coordinate on Ox).

Property3. For any radius vector $\overrightarrow{OB}$, the next is true: $x_B = OB \cdot \cos\alpha$ **[F]**

(where α is an angle of inclination of $\overrightarrow{OB}$).

Proof. If $\alpha = 0^o$, $\alpha = 90^o$, $\alpha = 180^o$, $\alpha = 270^o$,

then the equality **[F]** is obviously true.
Let we have any other case.
We fix any $\overrightarrow{OB}$ and draw the unit circle Ω with the center at O.
$\overrightarrow{OB}$ may not intersect Ω (if it's length is less than 1), in any case we draw the ray OB and mark the point C at which it intersects Ω **[pict13]**.



pict.13

Let's draw perpendiculars BB_x and CC_x. In any case the right triangles

ΔOBB_x and ΔOCC_x are similar, then $\dfrac{OB}{OC} = \dfrac{OB_x}{OC_x} \Rightarrow [as\ OC = 1] \Rightarrow OB_x = OB \cdot OC_x$ - from this equality follows the result we need (formula) **[F]**. We just have to consider all the possible cases for the inclination angle α.

[1] $\alpha \in (0^o, 90^o)$, then $x_B = OB_x$ and $\cos\alpha = OC_x$ and $OB_x = OB \cdot OC_x \Rightarrow x_B = OB \cdot \cos\alpha$.

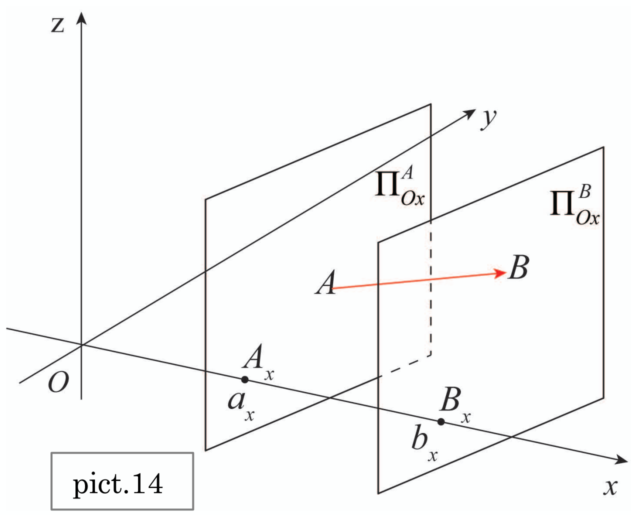
[2] $\alpha \in (90^o, 180^o)$, then $x_B = -OB_x$ and $\cos\alpha = -OC_x$ and
$OB_x = OB \cdot OC_x \Rightarrow -x_B = OB \cdot (-\cos\alpha) \Rightarrow x_B = OB \cdot \cos\alpha$.

[3] $\alpha \in (180^o, 270^o)$, then $x_B = -OB_x$ and $\cos\alpha = -OC_x$ and
$OB_x = OB \cdot OC_x \Rightarrow -x_B = OB \cdot (-\cos\alpha) \Rightarrow x_B = OB \cdot \cos\alpha$.

[4] $\alpha \in (270^o, 360^o)$, then $x_B = OB_x$ and $\cos\alpha = OC_x$ and $OB_x = OB \cdot OC_x \Rightarrow x_B = OB \cdot \cos\alpha$.

Vector coordinates in the space

Let $\overrightarrow{AB}$ is any vector in the space.
Let's draw the plane Π_{Ox}^{B} through the point B, which is perpendicular to Ox, it intersects the axis Ox at the point B_x with coordinate b_x **[pict14]**. Let's draw the plane Π_{Ox}^{A} through the point A, which is perpendicular to Ox, it intersects Ox at the point A_x with coordinate a_x. By definition, x-coordinate of $\overrightarrow{AB}$ is the difference $b_x - a_x$. In the exactly similar



pict.14

way we can define: y-coordinate of $\overrightarrow{AB}$ is the difference $b_y - a_y$. And z-coordinate of $\overrightarrow{AB}$ is $b_z - a_z$. And finally, coordinates of $\overrightarrow{AB}$ is the set of numbers $(b_x - a_x,\ b_y - a_y,\ b_z - a_z)$.

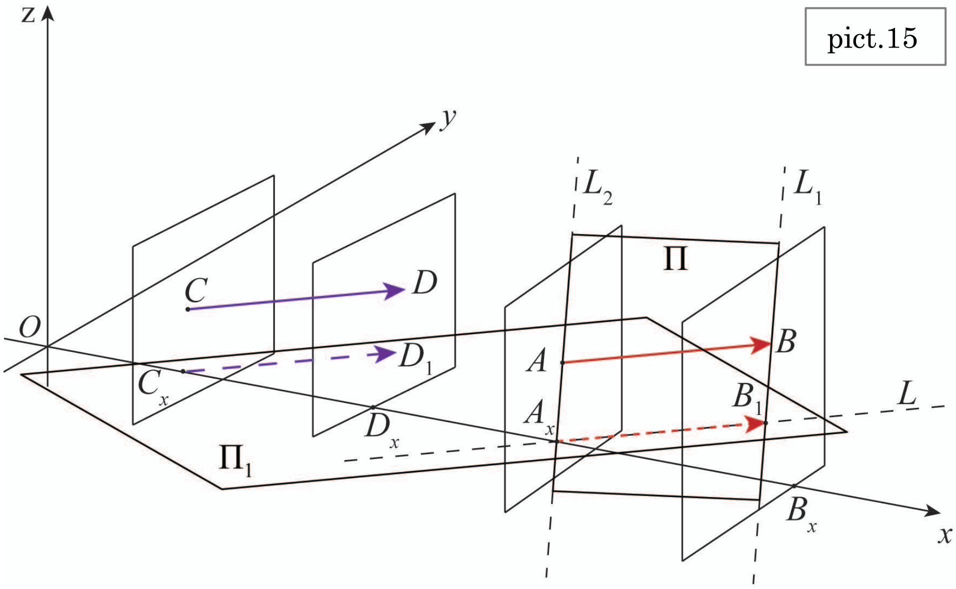
Theorem5. Equal vectors have equal coordinates.

Proof. [pict15] Let's show that equal vectors have equal x-coordinates.

Let's fix some vector $\overrightarrow{AB}$.

Let draw the planes Π_{Ox}^{B} and Π_{Ox}^{A}. The picture is done for the most general case, when Π_{Ox}^{B} and Π_{Ox}^{A} do not coincide, so these are different planes.

We can build the vector $\overrightarrow{A_x B_1}$



pict.15

such that $\overrightarrow{A_x B_1} = \overrightarrow{AB}$ (theorem4). Let's show that B_1 must belong to the plane Π_{Ox}^{B}.

[A] Let Π_{Ox}^{B} and Π_{Ox}^{A} coincide. As $\overrightarrow{A_x B_1}$ and $\overrightarrow{AB}$ are equal, they must be parallel $A_x B_1 \parallel AB$. Obviously A, B, A_x lie on the same plane $\Pi_{Ox}^{B} = \Pi_{Ox}^{A}$. If B_1 doesn't belong to $\Pi_{Ox}^{B} = \Pi_{Ox}^{A}$, then $A_x B_1, AB$ are skew lines and they are not parallel, which is not true. Then B_1 belongs to Π_{Ox}^{B}.

[B] Let Π_{Ox}^{B} and Π_{Ox}^{A} do not coincide. Then these are parallel planes (because they are both perpendicular to Ox). As $A_x B_1 \parallel AB$, we can draw the plane Π through the lines $A_x B_1$ and AB. The plane Π intersects Π_{Ox}^{B} and Π_{Ox}^{A} and there are some lines of intersection $\Pi \cap \Pi_{Ox}^{B} = L_1$ and $\Pi \cap \Pi_{Ox}^{A} = L_2$. There must be $L_1 \parallel L_2$ (it is one of the axioms of stereometry).

As L_2 passes through A, A_x, then $L_2 = AA_x$. And L_1 passes through B.

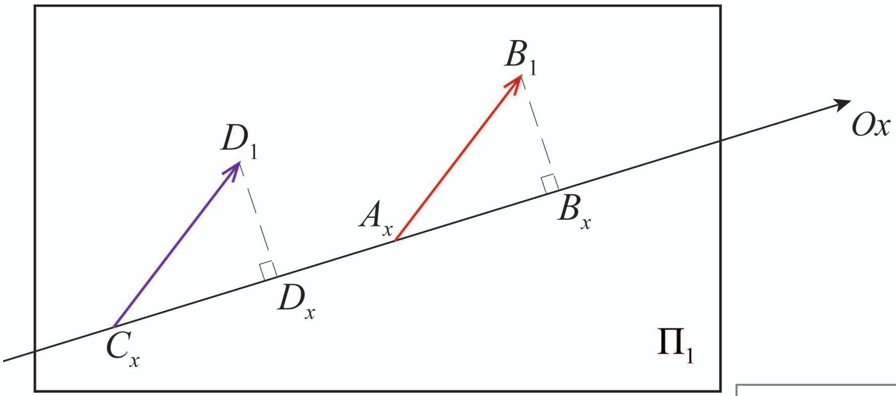
As $\overrightarrow{AB}, \overrightarrow{A_x B_1}$ are equal, lines AA_x and BB_1 must be parallel.

Let's sum up: $L_2 = AA_x$ is parallel to BB_1 and $L_2 = AA_x$ is parallel to $L_1 \subset \Pi_{Ox}^{B}$ which passes through B. Then (by transitivity) BB_1 must be parallel to L_1, both these lines pass through the point B, then they must coincide, therefore $B_1 \in \Pi_{Ox}^{B}$.

Next, $\overrightarrow{AB}, \overrightarrow{A_x B_1}$ obviously have equal x-coordinates (because their start points lie on Π_{Ox}^{A} and their end points lie on Π_{Ox}^{B}).

Let's take any other vector $\overrightarrow{CD}$ which is equal to $\overrightarrow{AB}$. In the same way we build $\overrightarrow{C_xD_1} = \overrightarrow{CD}$. And $\overrightarrow{C_xD_1}$ has the same x-coordinate as $\overrightarrow{CD}$.

By transitivity of the vector equality (theorem3) we have $\overrightarrow{A_xB_1} = \overrightarrow{C_xD_1}$, then A_xB_1 and C_xD_1 must be parallel. Then we can draw the plane Π_1 through A_xB_1 and C_xD_1, obviously Π_1 contains the axis Ox. Then we have the equal vectors $\overrightarrow{A_xB_1}, \overrightarrow{C_xD_1}$ and the axis Ox on the plane Π_1, the start points of these vectors lie on Ox **[pict15.1]**. Obviously $B_1B_x \perp Ox$ (because B_1B_x lies on the plane which is perpendicular to Ox) and also $D_1D_x \perp Ox$. Then x-coordinate of $\overrightarrow{A_xB_1}$ when we consider it as a space vector is the same as x-coordinate of $\overrightarrow{A_xB_1}$ when we consider it as a plane vector on Π_1. Obviously $\overrightarrow{A_xB_1}$ and $\overrightarrow{C_xD_1}$ are equal vectors on the plane Π_1, then their x-coordinates on Π_1 are equal: $b_x - a_x = d_x - c_x$ (property1), in the same time $b_x - a_x$ and $d_x - c_x$ are x-coordinates of the space vectors $\overrightarrow{A_xB_1}$ and $\overrightarrow{C_xD_1}$. Therefore x-coordinates of the space vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal. Similarly we can show that y-coordinates and z-coordinates of $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal. Then equal vectors have equal coordinates. Everything is proved.



pict.15.1

By definition, when we want to find coordinates of some point A (in the space) we need to draw the planes Π_x, Π_y, Π_z through point A which are perpendicular to Ox, Oy, Oz. These planes intersect axises at some points A_x, A_y, A_z with coordinates a_x, a_y, a_z. The numbers a_x, a_y, a_z are coordinates of A. Let's notice that $Ox \perp \Pi_x$ and therefore Ox is perpendicular to any line which lies on Π_x, in particular $Ox \perp AA_x$. Similarly $Oy \perp AA_y$ and $Oz \perp AA_z$, then AA_x, AA_y, AA_z are perpendiculars from A on Ox, Oy, Oz, and coordinates of any point A in the space can be found by drawing perpendiculars from A to coordinate axises.

Exercise. For any vector $\vec{a}$ with coordinates (a_x, a_y, a_z), the length of $\vec{a}$ is

$$a = \sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2}.$$

Solution: $\vec{a}$ connects it's start and end points, these points have some coordinates $(x_{\text{start}}, y_{\text{start}}, z_{\text{start}})$ and $(x_{\text{end}}, y_{\text{end}}, z_{\text{end}})$. The distance between these points is $\sqrt{(x_{\text{end}} - x_{\text{start}})^2 + (y_{\text{end}} - y_{\text{start}})^2 + (z_{\text{end}} - z_{\text{start}})^2}$ and here we see that

$x_{\text{end}} - x_{\text{start}}$, $y_{\text{end}} - y_{\text{start}}$, $z_{\text{end}} - z_{\text{start}}$ are coordinates of $\vec{a}$ (by definition), so these differences are a_x, a_y, a_z .

Def: a radius vector is any vector $\overrightarrow{OA}$, which start point O is located at the origin.

The start point O of $\overrightarrow{OA}$ has coordinates $(0,0,0)$, then: coordinates of any radius vector are exactly coordinates of it's end point A.

Theorem6. If two vectors have equal coordinates, then these vectors are equal.

Proof. Let $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are vectors with equal coordinates $(b_x - a_x,\ b_y - a_y,\ b_z - a_z) = (d_x - c_x,\ d_y - c_y,\ d_z - c_z) \equiv (a,b,c)$. Let's build the radius vector $\overrightarrow{OB_1} = \overrightarrow{AB}$. We have already shown that equal vectors have equal coordinates, therefore $\overrightarrow{OB_1}$ has coordinates (a,b,c). Then B_1 has coordinates (a,b,c).

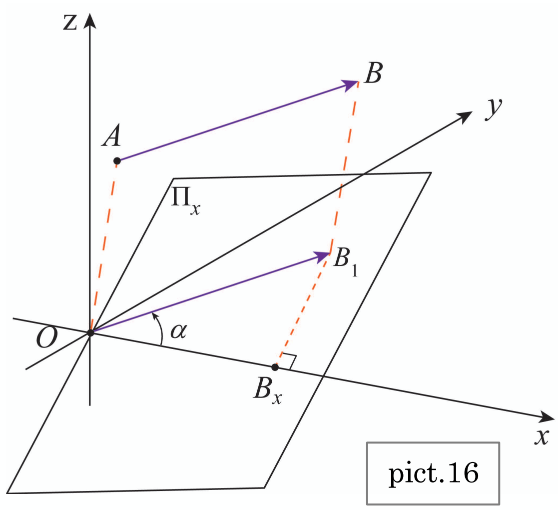
Let's build now the radius vector $\overrightarrow{OD_1} = \overrightarrow{CD}$. Equal vectors have equal coordinates, therefore $\overrightarrow{OD_1}$ has coordinates (a,b,c). Then D_1 has coordinates (a,b,c). We see that B_1 and D_1 have the same coordinates, therefore they must coincide: $B_1 = D_1$. Then $\overrightarrow{OB_1} = \overrightarrow{OD_1}$ and, by transitivity of the vector equality (**theorem3**), we get $\overrightarrow{AB} = \overrightarrow{CD}$.

Let's sum up: from the **theorems 5,6** follows

Theorem7. Two vectors are equal $\Leftrightarrow$ coordinates of these vectors are equal.

Another representation of vector coordinates.

Let $\overrightarrow{AB}$ is any vector in the space, it has coordinates $(b_x - a_x,\ b_y - a_y,\ b_z - a_z)$. Let's build the radius vector $\overrightarrow{OB_1} = \overrightarrow{AB}$ **[pict16]**. $\overrightarrow{OB_1}$ has the same coordinates: $(b_x - a_x,\ b_y - a_y,\ b_z - a_z)$. Then B_1 has coordinates $(b_x - a_x,\ b_y - a_y,\ b_z - a_z)$. Then, if we draw the perpendicular B_1B_x to Ox, the point B_x has x-coordinate $x_{B_1} = b_x - a_x$. Now we can consider $\overrightarrow{OB_1}$ as a plane vector on the plane Π_x which contains $\overrightarrow{OB_1}$ and Ox.



pict.16

The start point O of $\overrightarrow{OB_1}$ is located at the origin. According to the property3 (from above), we have: $x_{B_1} = OB_1 \cdot \cos\alpha \Leftrightarrow b_x - a_x = OB_1 \cdot \cos\alpha$ where $\alpha \in [0^o, 360^o)$ is the angle between Ox and $\overrightarrow{OB_1}$, which is counted in the counterclockwise direction from Ox.

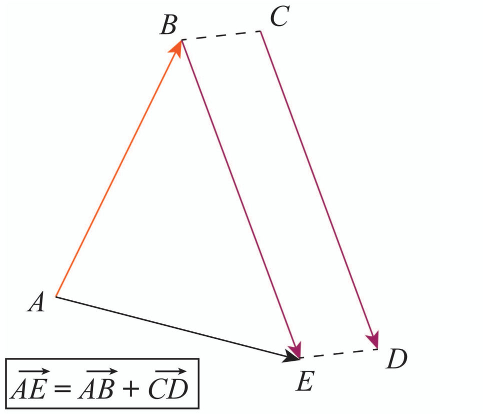
Then $b_x - a_x = OB_1 \cdot \cos\alpha \Rightarrow b_x - a_x = AB \cdot \cos\alpha$ and similarly $b_y - a_y = AB \cdot \cos\beta$ and $b_z - a_z = AB \cdot \cos\gamma$. Therefore, coordinates of $\overrightarrow{AB}$ are $(AB \cdot \cos\alpha, AB \cdot \cos\beta, AB \cdot \cos\gamma)$.

Def. We will say that $\overrightarrow{AB}$ forms the angle α with the axis Ox if the radius vector $\overrightarrow{OB_1} = \overrightarrow{AB}$ forms the angle α with the axis Ox. The angle α is called an inclination angle of $\overrightarrow{AB}$ with respect to Ox. And similarly, we define the angles of inclination β, γ with respect to Ox, Oy, Oz. By definition, $\cos\alpha, \cos\beta, \cos\gamma$ - "direction cosines" of $\overrightarrow{AB}$.

Then: if we multiply a module of a vector by it's directional cosines, we will get exactly coordinates of a vector. Such representation of vector coordinates is very important for problems of mechanics.

Addition of vectors

Def. The sum of vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ is the vector $\overrightarrow{AE} \equiv \overrightarrow{AB} + \overrightarrow{CD}$, which we get from the next process: we build $\overrightarrow{BE} = \overrightarrow{CD}$ **[pict17]**, then we build $\overrightarrow{AE}$, then $\overrightarrow{AE}$ is a sum of $\overrightarrow{AB}$ and $\overrightarrow{CD}$.



pict.17

Property4 From this definition immediately follows that: for any points A, B, C in the space we have: $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$

Assertion3. When we add vectors, their coordinates must be added. For any vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$: x-coordinate/ y-coordinate/ z-coordinate of $\overrightarrow{AB} + \overrightarrow{CD}$ is a sum of x-coordinates/ y-coordinates/ z-coordinates of vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$.

Proof. $\overrightarrow{AB}$ has coordinates $(b_x - a_x,\ b_y - a_y,\ b_z - a_z)$ and $\overrightarrow{CD}$ has coordinates $(\theta_1,\ \theta_2,\ \theta_3)$. Let's find the sum $\overrightarrow{AB} + \overrightarrow{CD}$, we need to build $\overrightarrow{BE} = \overrightarrow{CD}$ **[pict17]**. As $\overrightarrow{BE}$ connects B and E, it has coordinates $(e_x - b_x,\ e_y - b_y,\ e_z - b_z)$. On the other hand $\overrightarrow{BE}$ has the same coordinates as $\overrightarrow{CD}$ (because these are equal vectors), then $(e_x - b_x,\ e_y - b_y,\ e_z - b_z) = (\theta_1,\ \theta_2,\ \theta_3)$ **[J]**.

By definition, the sum of our vectors is the vector $\overrightarrow{AE}$ which connects A and E, so it has coordinates $(e_x - a_x,\ e_y - a_y,\ e_z - a_z)$. We obviously have:

$$(e_x - a_x,\ e_y - a_y,\ e_z - a_z) = ((e_x - b_x) + (b_x - a_x),\ (e_y - b_y) + (b_y - a_y),\ (e_z - b_z) + (b_z - a_z)),$$

then, by **[J]**, we can rewrite:

$(e_x - a_x,\ e_y - a_y,\ e_z - a_z) = ((e_x - b_x) + \theta_1,\ (e_y - b_y) + \theta_2,\ (e_z - b_z) + \theta_3)$. Everything is proved.

Sometimes when we speak about vectors we do not have any necessity to specify their start and end points, especially when we speak about general properties of vectors, or consider some collections of vectors, in such case we denote vectors just like $\vec{a},\vec{b},\vec{c}...$, it's very convenient, because such notation takes less place.

Assertion3 can be rewritten as: for any vectors $\vec{a}$ and $\vec{b}$ with coordinates (a_x,a_y,a_z) and (b_x,b_y,b_z), their sum $\vec{a}+\vec{b}$ has coordinates $(a_x+b_x,a_y+b_y,a_z+b_z)$.

Theorem8. Addition of vectors is commutative: $\vec{a}+\vec{b}=\vec{b}+\vec{a}$ (for any vectors $\vec{a},\vec{b}$) and associative $(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})$ (for any vectors $\vec{a},\vec{b},\vec{c}$).

Let's prove commutativity. Let (a_x,a_y,a_z) and (b_x,b_y,b_z) are coordinates of $\vec{a}$ and $\vec{b}$.
Then $\vec{a}+\vec{b}$ has coordinates $(a_x+b_x,a_y+b_y,a_z+b_z)$, and $\vec{b}+\vec{a}$ has coordinates, $(b_x+a_x,b_y+a_y,b_z+a_z)$. We see that vectors $\vec{a}+\vec{b}$ and $\vec{b}+\vec{a}$ have equal coordinates. then, according to the theorem7, $\vec{a}+\vec{b}$ and $\vec{b}+\vec{a}$ are equal.

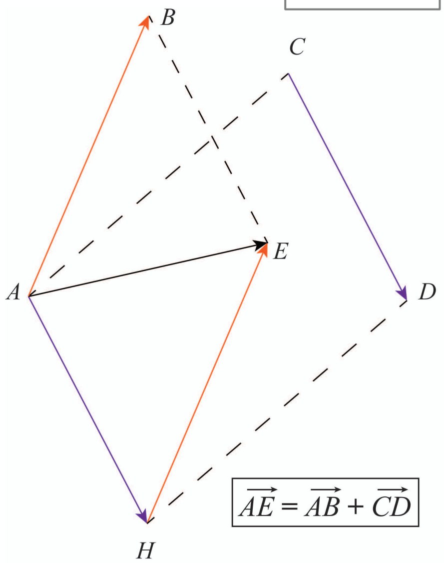
Next, associativity. **[1]** $\vec{a}+\vec{b}$ has coordinates $(a_x+b_x,a_y+b_y,a_z+b_z)$, then $(\vec{a}+\vec{b})+\vec{c}$ has coordinates $((a_x+b_x)+c_x,\ (a_y+b_y)+c_y,\ (a_z+b_z)+c_z)$.

[2] $\vec{a}$ has coordinates (a_x,a_y,a_z), and $(\vec{b}+\vec{c})$ has coordinates $(b_x+c_x,\ b_y+c_y,\ b_z+c_z)$.

Then $\vec{a}+(\vec{b}+\vec{c})$ has coordinates $(a_x+(b_x+c_x),\ a_y+(b_y+c_y),\ a_z+(b_z+c_z))$.

We see that $(\vec{a}+\vec{b})+\vec{c}$ and $\vec{a}+(\vec{b}+\vec{c})$ have equal coordinates, then (theorem6) these vectors are equal.

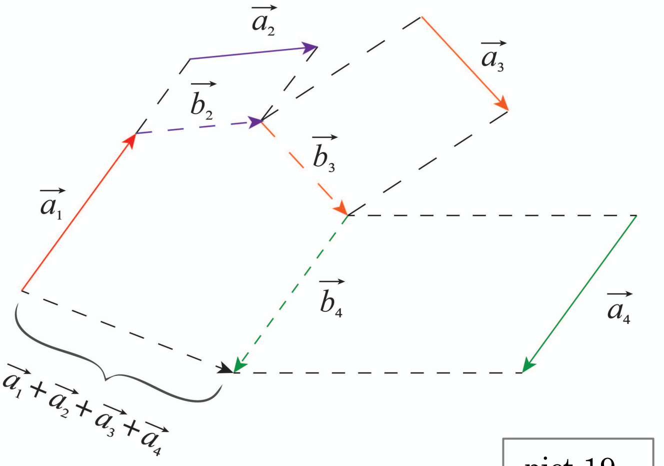
pict.18



Consequence1. Sums of equal vectors are equal vectors:
$\vec{a}_1=\vec{b}_1,\vec{a}_2=\vec{b}_2$ .... $\vec{a}_n=\vec{b}_n$, then $\vec{a}_1+\vec{a}_2+...+\vec{a}_n=\vec{b}_1+\vec{b}_2+...+\vec{b}_n$ (really, as vectors $\vec{a}_1+\vec{a}_2+...+\vec{a}_n$ and $\vec{b}_1+\vec{b}_2+...+\vec{b}_n$ have equal coordinates, then these vectors are equal).

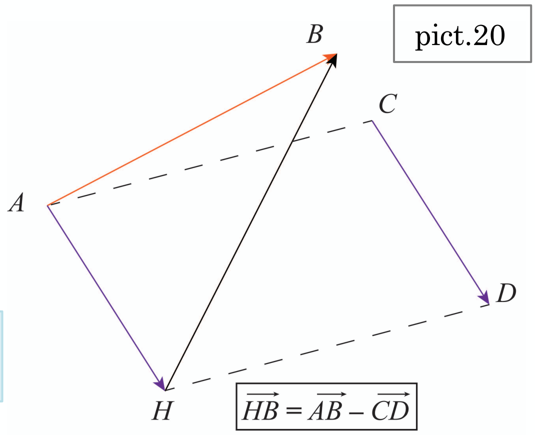
Parallelogram rule. For any two not collinear vectors $\overrightarrow{AB},\overrightarrow{CD}$ the sum of these vectors can be found in the next way **[pict18]**: we build $\overrightarrow{AH}=\overrightarrow{CD}$ and we draw the parallelogram, which adjacent sides are AB and AH, then the diagonal vector $\overrightarrow{AE}$ is a sum of vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$.

Chain rule [**pict19**]. If we want to find the sum of several vectors $\vec{a}_1, \vec{a}_2, \dots\ \vec{a}_n$ **[CH]**, we can construct the "chain" from these vectors: the end point of $\vec{a}_1$ must be taken as a start point of $\vec{b}_2 = \vec{a}_2$, the end point of $\vec{b}_2$ must be taken as a start point of $\vec{b}_3 = \vec{a}_3$ ..., the end point of $\vec{b}_{n-1} = \vec{a}_{n-1}$ must be taken as a start point of $\vec{b}_n = \vec{a}_n$. The vector $\vec{\Sigma}$, which connects the start point of $\vec{a}_1$ with the end point of $\vec{b}_n$, is a sum of all vectors **[CH]**: $\vec{\Sigma} = \vec{a}_1 + \vec{a}_2 + \dots + \vec{a}_n$.



pict.19

Subtraction of vectors. The difference of vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ is the vector $\overrightarrow{HB} = \overrightarrow{AB} - \overrightarrow{CD}$ [**pict20**], which we get from the next process: we build $\overrightarrow{AH} = \overrightarrow{CD}$, then we build $\overrightarrow{HB}$, then $\overrightarrow{HB}$ is a difference of $\overrightarrow{AB}$ and $\overrightarrow{CD}$.



pict.20

Property5. $\overrightarrow{HB} = \overrightarrow{AB} - \overrightarrow{CD} \Rightarrow \overrightarrow{HB} + \overrightarrow{CD} = \overrightarrow{AB}$.

Proof. $\Rightarrow$We have $\overrightarrow{HB} = \overrightarrow{AB} - \overrightarrow{CD}$, let's consider ΔAHB (property3) for any points A, H, B in the space we have:

$$\overrightarrow{AH} + \overrightarrow{HB} = \overrightarrow{AB} \Rightarrow /\,commutativity\,/ \Rightarrow \overrightarrow{HB} + \overrightarrow{AH} = \overrightarrow{AB} \Rightarrow$$

$$\Rightarrow [consequence\ 1:\ \ \overrightarrow{HB} = \overrightarrow{HB},\ \ \overrightarrow{AH} = \overrightarrow{CD}] \Rightarrow \overrightarrow{HB} + \overrightarrow{CD} = \overrightarrow{AB}.$$

Theorem9. When we subtract vectors, their coordinates must be subtracted, i.e., for any vectors $\vec{a}, \vec{b}$ with coordinates (a_x, a_y, a_z) and (b_x, b_y, b_z), their difference $\vec{a} - \vec{b}$ has coordinates $(a_x - b_x, a_y - b_y, a_z - b_z)$.

Proof. Let $\vec{a} - \vec{b} \equiv \vec{c}$ and $\vec{c}$ has some coordinates (c_x, c_y, c_z). From the property5 follows that $\vec{a} - \vec{b} \equiv \vec{c} \Rightarrow \vec{a} = \vec{c} + \vec{b}$. We have proved that when we add vectors, their coordinates must be added (assertion3): $(a_x, a_y, a_z) = (c_x + b_x, c_y + b_y, c_z + b_z) \Rightarrow a_x = c_x + b_x, a_y = c_y + b_y, a_z = c_z + b_z \Rightarrow$ $c_x = a_x - b_x, c_y = a_y - b_y, c_z = a_z - b_z$, then $(c_x, c_y, c_z) = (a_x - b_x, a_y - b_y, a_z - b_z)$.

Everything is proved.

Theorem10. $\vec{c} = \vec{a} - \vec{b} \Leftrightarrow \vec{c} + \vec{b} = \vec{a}$.

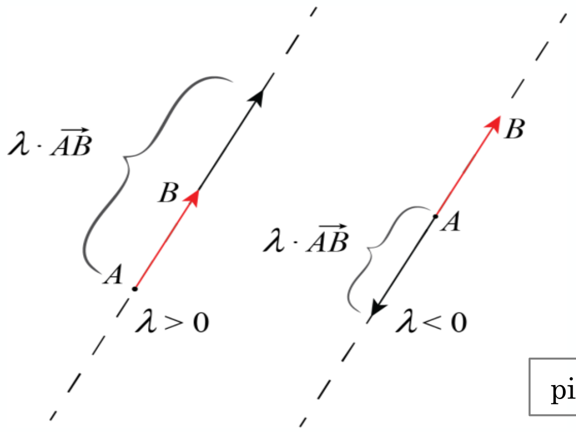
Proof. The part $\vec{c} = \vec{a} - \vec{b} \Rightarrow \vec{c} + \vec{b} = \vec{a}$ is proved (property5).

Conversely. Let $\vec{c} + \vec{b} = \vec{a}$, we want to show that $\vec{c} = \vec{a} - \vec{b}$. From $\vec{c} + \vec{b} = \vec{a}$ follows that coordinates of $\vec{a}$ are sums of coordinates of $\vec{c}$ and $\vec{b}$, so

$(a_x, a_y, a_z) = (c_x + b_x, c_y + b_y, c_z + b_z) \Rightarrow a_x = c_x + b_x, a_y = c_y + b_y, a_z = c_z + b_z \Rightarrow$
$\Rightarrow c_x = a_x - b_x, c_y = a_y - b_y, c_z = a_z - b_z$ **[M]**. According to the theorem9, $\vec{a} - \vec{b}$ has coordinates $(a_x - b_x, a_y - b_y, a_z - b_z)$, from **[M]** we see that $\vec{c}$ has the same coordinates as $\vec{a} - \vec{b}$,

Then (theorem7) vectors $\vec{c}$ and $\vec{a} - \vec{b}$ are equal: $\vec{c} = \vec{a} - \vec{b}$.

Multiplication by numbers [pict21].



pict.21

$\overrightarrow{AB}$ is any vector. We define

$0 \cdot \overrightarrow{AB} \equiv /\mathit{by\ definition}/ \equiv \overrightarrow{AA} = \vec{0}$

For any positive number $\lambda > 0$ we define:

$\lambda \cdot \overrightarrow{AB}$ is the vector which start point is A, which is co-directed with $\overrightarrow{AB}$, and which's length is $\lambda \cdot AB$.

For any negative number $\lambda < 0$ we define: $\lambda \cdot \overrightarrow{AB}$ is the vector which start point is A, which is anti-directed with $\overrightarrow{AB}$, and which's length is $|\lambda| \cdot AB$.

Def. $\vec{a}$ and $\vec{b}$ are non-zero vectors, $\vec{a}$ is called proportional to $\vec{b}$ if there exist some constant $\lambda \in R$ such that $\vec{a} = \lambda \cdot \vec{b}$

Exercise. Show that if $\vec{a}$ is proportional to $\vec{b}$, then $\vec{b}$ is proportional to $\vec{a}$.

From here follows: if one of non-zero vectors $\vec{a}, \vec{b}$ is proportional to the other, then we can just say “these vectors are proportional”.

Exercise. Non-zero vectors $\vec{a}, \vec{b}$ are proportional $\Leftrightarrow$ these vectors lie on parallel lines (in particular, on the same line).

Theorem11. If some vector is multiplied by λ, coordinates of that vector must be also multiplied by λ.

Proof. Let's take any vector $\overrightarrow{AB}$, it has coordinates (Another representation of vector coordinates) $(AB \cdot \cos\alpha, AB \cdot \cos\beta, AB \cdot \cos\gamma)$. Let λ is positive, then $\lambda \cdot \overrightarrow{AB}$ forms exactly the same angles

α, β, γ with coordinate axises, and it's length (by definition) is $\lambda \cdot AB$, then $\lambda \cdot \overrightarrow{AB}$ has coordinates $((\lambda \cdot AB) \cdot \cos\alpha,\ (\lambda \cdot AB) \cdot \cos\beta,\ (\lambda \cdot AB) \cdot \cos\gamma)$.

We see that coordinates of $\lambda \cdot \overrightarrow{AB}$ are exactly coordinates of $\overrightarrow{AB}$, which are multiplied by λ.
When $\lambda = 0$ the **theorem11** is obviously true.

Let now $\lambda < 0$. So $\overrightarrow{AB}$ forms some angles α, β, γ with Ox, Oy, Oz, then $\lambda \cdot \overrightarrow{AB} \| \lambda < 0$ (which direction is opposite to $\overrightarrow{AB}$) forms the angles $(180^o \pm \alpha),\ (180^o \pm \beta),\ (180^o \pm \gamma)$ with Ox, Oy, Oz, and it's length is $|\lambda| \cdot AB$, then it has coordinates

$$(|\lambda| \cdot AB \cdot \cos(180^o \pm \alpha),\ |\lambda| \cdot AB \cdot \cos(180^o \pm \beta),\ |\lambda| \cdot AB \cdot \cos(180^o \pm \gamma)) =$$

$$= (-|\lambda| \cdot AB \cdot \cos\alpha,\ -|\lambda| \cdot AB \cdot \cos\beta,\ -|\lambda| \cdot AB \cdot \cos\gamma) = [as\ \lambda = -|\lambda|] =$$

$= (\lambda \cdot AB \cdot \cos\alpha,\ \lambda \cdot AB \cdot \cos\beta,\ \lambda \cdot AB \cdot \cos\gamma)$ and we see again that coordinates of $\lambda \cdot \overrightarrow{AB}$ are exactly coordinates of $\overrightarrow{AB}$, which are multiplied by λ.

Def. For any vector $\vec{a}$ we define $-\vec{a} \equiv (-1) \cdot \vec{a}$.

Exercise. For any vectors $\vec{a}, \vec{b}$ we have: **[A]** $\vec{a} + (-\vec{a}) = \vec{a} - \vec{a} = \vec{0}$ and **[B]** $\vec{a} - \vec{b} = \vec{a} + (-\vec{b})$ **[C]** $\vec{a} - (-\vec{b}) = \vec{a} + \vec{b}$. **[D]** $(\lambda \cdot \mu) \cdot \vec{a} = \lambda \cdot (\mu \cdot \vec{a})$ (for any $\lambda, \mu \in R$), **[E]** $\lambda \cdot \vec{a} + \mu \cdot \vec{a} = (\lambda + \mu) \cdot \vec{a}$ (for any $\lambda, \mu \in R$), **[F]** $\lambda \cdot \vec{a} + \lambda \cdot \vec{b} = \lambda \cdot (\vec{a} + \vec{b})$ (for any $\lambda \in R$).

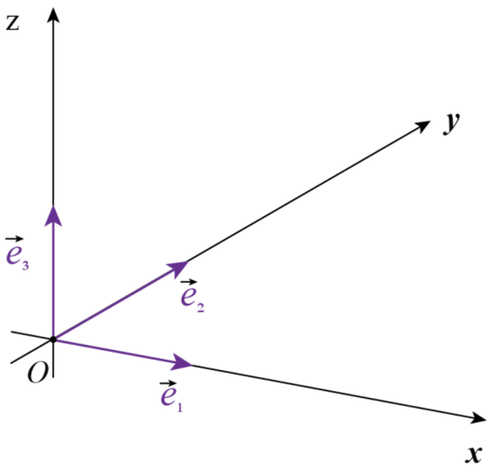
Hint. In each case it's very easy to show that vectors on the left and on the right part of equality have equal coordinates. And if vectors have equal coordinates, then these vectors are equal (**theorem7**).

Def. Let $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ are some vectors and $\lambda_1, \lambda_2 \ldots \lambda_n$ are some real numbers.
The vector $\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 \ldots + \lambda_n \vec{a}_n$ is called a linear combination of $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ with coefficients $\lambda_1, \lambda_2 \ldots \lambda_n$, or just a linear combination of $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$.

Property5. Obviously, x-coordinate of $\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 \ldots + \lambda_n \vec{a}_n$ is a linear combination of x-coordinates of vectors $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ with coefficients $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$.
And y-coordinate of $\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 \ldots + \lambda_n \vec{a}_n$ is a linear combination of y-coordinates of vectors $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ with coefficients $\lambda_1, \lambda_2 \ldots \lambda_n$. And the same is true for z-coordinate .

Basis

Def. Let's take the "unit vectors" $\vec{e}_x, \vec{e}_y, \vec{e}_z$ with coordinates $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ **[pict22]**.
The start point of each vector is located at the origin O, these vectors are called the standard basis vectors.



pict.22

Any vector $\vec{a}$ with coordinates (a_x, a_y, a_z) can be represented as a linear combination of these vectors:
$\vec{a} = a_x\vec{e}_x + a_y\vec{e}_y + a_z\vec{e}_z$ (property5) and this representation is unique. Really, let $\vec{a} = \lambda\vec{e}_x + \mu\vec{e}_y + \delta\vec{e}_z$ (where λ, μ, δ are unknown variables). As vectors $\vec{a}$, $\lambda\vec{e}_x + \mu\vec{e}_y + \delta\vec{e}_z$ are equal, they must have equal coordinates.
$\vec{a}$ has coordinates (a_x, a_y, a_z) and $\lambda\vec{e}_x + \mu\vec{e}_y + \delta\vec{e}_z$ has coordinates
$\lambda\cdot(1,0,0) + \mu\cdot(0,1,0) + \delta\cdot(0,0,1) = (\lambda,0,0) + (0,\mu,0) + (0,0,\delta) = (\lambda,\mu,\delta)$, and there must be $(a_x, a_y, a_z) = (\lambda,\mu,\delta)$, then $a_x = \lambda$, $a_y = \mu$, $a_z = \delta$ and the representation is unique.

Projection on an axis. Let's fix an arbitrary axis L in the space.
(**Notice**: when we fix some axis in the space, the scale on it must be the same as a scale on Ox, Oy, Oz).
Let $\overrightarrow{AB}$ is any vector. Let's draw two planes Π_L^B and Π_L^A through the points B and A, which are perpendicular to L, they intersect the axis at some points with coordinates x_B^L and x_A^L, the difference $x_B^L - x_A^L$ is called a projection of $\overrightarrow{AB}$ on the axis L. It's easy to notice now that x-coordinate of any vector $\vec{a}$ is just a projection of $\vec{a}$ on the axis Ox, and the same is true for y, z-coordinates of any vector. In fact, any axis in the space can be chosen from the very beginning as a coordinate axis Ox (for example). And we can get the same results, as we got earlier for Ox, for any axis L in the space. Let O_L is a point with a zero coordinate on L, let's build $\overrightarrow{O_LB_1} = \overrightarrow{AB}$.
By definition, the angle between $\overrightarrow{AB}$ and L is the angle $\alpha \in [0^o, 360^o)$ between $\overrightarrow{O_LB_1}$ and L, which is counted from L in the counterclockwise direction.
Then $AB\cdot\cos\alpha = x_B^L - x_A^L$ (the exactly similar result, page 17).

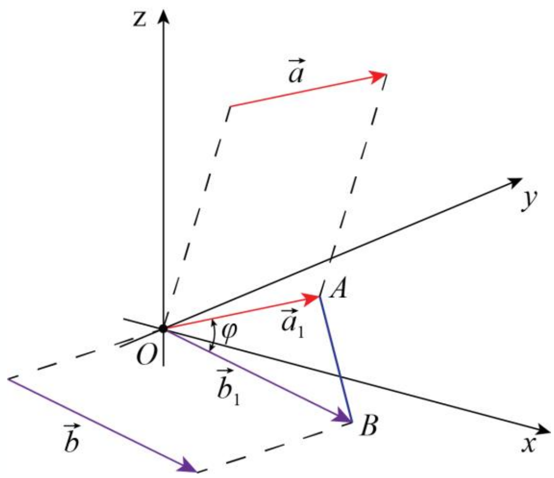
Theorem12. For any vectors $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ and any axis L, the projection of a sum-vector $\vec{a}_1 + \vec{a}_2 + \ldots + \vec{a}_n$ on L is equal to the sum of projections of $\vec{a}_1, \vec{a}_2 \ldots \vec{a}_n$ on L.

Comment: this theorem is exactly similar to the assertion: x-coordinate of a sum of vectors is a sum of x-coordinates of these vectors.

Dot product. For any vectors $\vec{a}$ and $\vec{b}$ with coordinates (a_x, a_y, a_z) and (b_x, b_y, b_z), the number $a_x b_x + a_y b_y + a_z b_z$ is called "a dot product of vectors $\vec{a}$ and $\vec{b}$", and we denote it $\vec{a} \cdot \vec{b} \equiv a_x b_x + a_y b_y + a_z b_z$.

Def: Let $\vec{a}$ and $\vec{b}$ are some non-zero vectors in the space. Let's draw the radius vectors $\vec{b}_1 = \vec{b}$ and $\vec{a}_1 = \vec{a}$ **[pict23]**. By definition, the angle between $\vec{a}$ and $\vec{b}$ is the angle $\varphi \in [0^o, 180^o]$ between $\vec{a}_1$ and $\vec{b}_1$.

Assertion4. The dot product of any non-zero vectors $\vec{a}, \vec{b}$ can be calculated as: $\vec{a} \cdot \vec{b} = a \cdot b \cdot \cos\varphi$, where φ is an angle between $\vec{a}$ and $\vec{b}$.



pict.23

Proof. Let $\vec{a}$ and $\vec{b}$ are vectors with coordinates (a_x, a_y, a_z) and (b_x, b_y, b_z), we build $\vec{b}_1 = \vec{b}$ and $\vec{a}_1 = \vec{a}$ **[pict23]**.

Let's denote $\vec{b}_1 = \overrightarrow{OB}$ and $\vec{a}_1 = \overrightarrow{OA}$. Then $\overrightarrow{OA}$ has coordinates (a_x, a_y, a_z) and $\overrightarrow{OB}$ has coordinates (b_x, b_y, b_z).

As the origin O has coordinates $(0,0,0)$, then A has coordinates (a_x, a_y, a_z) and B has coordinates (b_x, b_y, b_z).

Let's consider ΔOAB, here $OA = a$ and $OB = b$ and $\angle AOB = \varphi$.

Let's simplify the expression $a \cdot b \cdot \cos\varphi$ by using the cosine law for ΔOAB:

$$a^2 + b^2 - 2ab\cos\varphi = AB^2 \Rightarrow \cos\varphi = \frac{a^2 + b^2 - AB^2}{2ab}, \text{ then}$$

$$ab \cdot \cos\varphi = ab \cdot \frac{a^2 + b^2 - AB^2}{2ab} = \frac{a^2 + b^2 - AB^2}{2} =$$

$$= \frac{\left(\sqrt{a_x^2 + a_y^2 + a_z^2}\right)^2 + \left(\sqrt{b_x^2 + b_y^2 + b_z^2}\right)^2 - \left(\sqrt{(b_x - a_x)^2 + (b_y - a_y)^2 + (b_z - a_z)^2}\right)^2}{2} =$$

$$= \frac{a_x^2 + a_y^2 + a_z^2 + b_x^2 + b_y^2 + b_z^2 - \left((b_x - a_x)^2 + (b_y - a_y)^2 + (b_z - a_z)^2\right)}{2} =$$

$= a_x b_x + a_y b_y + a_z b_z = /by\ definition/ = \vec{a} \cdot \vec{b}$. Everything is proved.

The assertion4 allows us to find an angle between any non-zero vectors $\vec{a}$ and $\vec{b}$, if we know coordinates of these vectors.

For any (non-zero) $\vec{a}$ and $\vec{b}$, we can fix an arbitrary point D in the space and build $\overrightarrow{DA} = \vec{a}$ and $\overrightarrow{DB} = \vec{b}$, the angle $\varphi \in [0^o, 180^o]$ between $\overrightarrow{DA}$ and $\overrightarrow{DB}$ is obviously equal to the angle between

$\vec{a}$ and $\vec{b}$ (because the triangle, which is built on the radius vectors $\vec{a}_1 = \overrightarrow{OA}$ and $\vec{b}_1 = \overrightarrow{OB}$, is equal to the triangle which is built on $\overrightarrow{DA}$ and $\overrightarrow{DB}$).

Exercise1. Find the angle φ between vectors $\vec{a}$ and $\vec{b}$ with coordinates $(1,2,3)$ and $(4,5,6)$.

Solution. Let's write the dot product of these vectors in two different ways.

$$1\cdot4+2\cdot5+3\cdot6=\sqrt{1^2+2^2+3^2}\cdot\sqrt{4^2+5^2+6^2}\cdot\cos\varphi \Rightarrow \cos\varphi=\frac{32}{\sqrt{14}\cdot\sqrt{77}} \Rightarrow$$

$$\Rightarrow \cos\varphi=\frac{32}{7\sqrt{2}\cdot\sqrt{11}}=\frac{32}{7\sqrt{22}}.$$

Now we have the cosine value of the angle $\varphi\in[0^o,180^o]$. Any angle from $[0^o,180^o]$ is uniquely defined by it's cosine. So, our answer is: φ is such angle from $[0^o,180^o]$ that $\cos\varphi=\frac{32}{7\sqrt{22}}$.

Comment: we haven't defined the inverse trigonometric functions yet, but anyway, the answer "φ is arccosine of something" is not better than our answer.

Exercise2. Find the area of the triangle ΔABC which adjacent sides are the vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$ with coordinates $(1,2,3)$ and $(4,5,6)$.

Solution. We have already found the angle between these vectors. In order to find the area we can use the formula $S=\frac{1}{2}ab\sin\alpha$. So $AB=\sqrt{1^2+2^2+3^2}=\sqrt{14}$ and $AC=\sqrt{4^2+5^2+6^2}=\sqrt{77}$, the angle φ between these sides is such that $\cos\varphi=\frac{32}{7\sqrt{22}}$ and $\varphi\in[0^o,180^o]$, then

$$\sin\varphi=\sqrt{1-\cos^2\varphi}=\sqrt{1-\left(\frac{32}{7\sqrt{22}}\right)^2}=\sqrt{\frac{49\cdot22-32\cdot32}{49\cdot22}}=\frac{\sqrt{54}}{7\sqrt{22}}.$$

Then the area equals $S=\frac{1}{2}\sqrt{14}\cdot\sqrt{77}\cdot\frac{\sqrt{54}}{7\sqrt{22}}=\frac{1}{2}\sqrt{54}=\frac{3}{2}\sqrt{6}$.

Properties of a dot product.

[1] Non-zero vectors $\vec{a}$ and $\vec{b}$ are perpendicular if and only if their dot product is zero.

Proof. As we know $\vec{a}\cdot\vec{b}=a\cdot b\cdot\cos\varphi$. If vectors are perpendicular $\vec{a}\perp\vec{b}$, then $\varphi=90^o$, then $\vec{a}\cdot\vec{b}=a\cdot b\cdot\cos90^o=a\cdot b\cdot0=0$. Conversely, if the dot product is zero $\vec{a}\cdot\vec{b}=0$, then $a\cdot b\cdot\cos\varphi=0$. Vectors $\vec{a},\vec{b}$ are non-zero vectors, it means that $a\neq0, b\neq0$, then from the equality $a\cdot b\cdot\cos\varphi=0$ follows that $\cos\alpha=0\Rightarrow\alpha=90^o$.

By using **[1]** we can understand very quickly, are given vectors perpendicular or not. For example, vectors with coordinates $(1,2,3)$ and $(-3,-2,-1)$ are not perpendicular, because $1\cdot(-3)+2\cdot(-2)+3\cdot(-1)=-10$. And vectors with coordinates $(1,2,1)$ and $(-1,1,-1)$ are perpendicular, because $1\cdot(-1)+2\cdot1+1\cdot(-1)=0$.

[2] The dot product has a **linear property**:

[A] For any number λ and any vectors $\vec{a},\vec{b}$ we have: $(\lambda\vec{a})\cdot\vec{b}=\vec{a}\cdot(\lambda\vec{b})=\lambda\cdot(\vec{a}\cdot\vec{b})$

[B] For any vectors $\vec{a},\vec{b},\vec{c}$ we have: $(\vec{a}+\vec{b})\cdot\vec{c}=\vec{a}\cdot\vec{c}+\vec{b}\cdot\vec{c}$ and $\vec{c}\cdot(\vec{a}+\vec{b})=\vec{c}\cdot\vec{a}+\vec{c}\cdot\vec{b}$.

Proof. Both proofs are very simple and immediately follow from the definition of a dot product: $\vec{a}\cdot\vec{b}\equiv a_xb_x+a_yb_y+a_zb_z$. Let's denote coordinates of our vectors:

$\vec{a}=(a_x,a_y,a_z)$, $\vec{b}=(b_x,b_y,b_z)$, $\vec{c}=(c_x,c_y,c_z)$ then, for example, $\vec{a}+\vec{b}$ has coordinates $(a_x+b_x,a_y+b_y,a_z+b_z)$ and the dot product

$$(\vec{a}+\vec{b})\cdot\vec{c}=(a_x+b_x)\cdot c_x+(a_y+b_y)\cdot c_y+(a_z+b_z)\cdot c_z=$$
$$=\left[a_x\cdot c_x+a_y\cdot c_y+a_z\cdot c_z\right]+\left[b_x\cdot c_x+b_y\cdot c_y+b_z\cdot c_z\right]=\vec{a}\cdot\vec{c}+\vec{b}\cdot\vec{c}$$ and etc.

Change of approach. In mathematics equal vectors are usually understanded as a same vector.
So usually people point on equal vectors (which do not actually coincide) and say that these vectors are a same vector. It is only a mathematical approach, and it obviously doesn't work in mechanics. Really, suppose that several forces exert on some object. Then each force is unique and the point of application of each force is very important. If some two forces are equal as vectors, but they have different points of application, these are different forces anyway, it's not right to understand these forces as a same force.
But in mathematics equal vectors are usually perceived as a same vector.

Most important trigonometric formulas

Remark1. For any angle α and for any integer number k: $\cos(360^0\cdot k+\alpha)=\cos\alpha$ and $\sin(360^0\cdot k+\alpha)=\sin\alpha$.

Remark2. We have shown (Book I, page170) that for any angle γ **[1]** $\sin(-\gamma)=-\sin(\gamma)$ and **[2]** $\cos(-\gamma)=\cos\gamma$. Therefore, for any angles α,β we can take $\gamma=\beta-\alpha$ and then $\sin(\alpha-\beta)=-\sin(\beta-\alpha)$ and $\cos(\alpha-\beta)=\cos(\beta-\alpha)$.

Assertion5. For any angles α and β: $\cos(\alpha-\beta)=\cos\alpha\cdot\cos\beta+\sin\alpha\cdot\sin\beta$ **[A]**

Proof. Each angle α and β can be represented as: $\alpha=360^o\cdot k+\alpha_1$, here k is an integer number and $\alpha_1\in[0^o,360^o)$ and $\beta=360^o\cdot m+\beta_1$ where m is an integer number and $\beta_1\in[0^o,360^o)$.

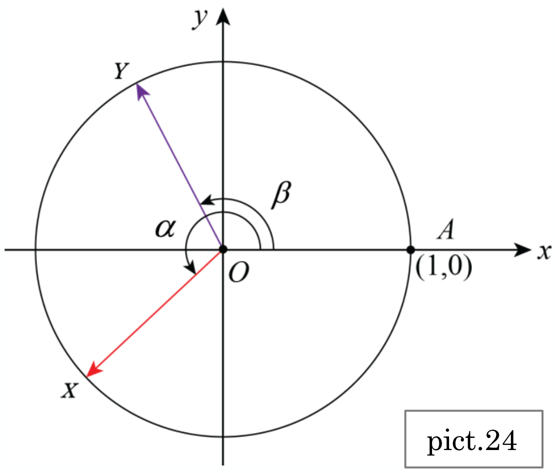
Let's simplify both sides of **[P]** separately:

$$\cos(\alpha+\beta)=\cos(360^o\cdot k+\alpha_1-(360^o\cdot m+\beta_1))=\cos(360^o\cdot(k-m)+(\alpha_1-\beta_1))=$$
$$=//\,remark\ \ 1//=\cos(\alpha_1-\beta_1).$$

$$\cos\alpha\cdot\cos\beta+\sin\alpha\cdot\sin\beta=\cos(360^o\cdot k+\alpha_1)\cdot\cos(360^o\cdot m+\beta_1)+\sin(360^o\cdot k+\alpha_1)\cdot\sin(360^o\cdot m+\beta_1)=$$
$=//\,remark1//=\cos\alpha_1\cdot\cos\beta_1+\sin\alpha_1\cdot\sin\beta_1$. Therefore it is enough to prove that

$\cos(\alpha_1-\beta_1)=\cos\alpha_1\cdot\cos\beta_1+\sin\alpha_1\cdot\sin\beta_1$ for $\alpha_1\in[0^o,360^o)$ and $\beta_1\in[0^o,360^o)$.

We can discard indexes here, so we have to prove the equality **[A]** only for $\alpha, \beta \in [0^o, 360^o)$, let's fix any $\alpha, \beta \in [0^o, 360^o)$ and let's build the angles $\angle AOX = \alpha$ and $\angle AOY = \beta$ **[pict24]**. We have two unit vectors $\overrightarrow{OX}$ and $\overrightarrow{OY}$, the angle between these vectors is equal to $(\alpha - \beta)$ (if $\alpha \geq \beta$) or $(\beta - \alpha)$ (if $\beta \geq \alpha$) in any case, the dot product of these vectors



pict.24

$\overrightarrow{OX} \cdot \overrightarrow{OY} = OX \cdot OY \cdot \cos(\alpha - \beta)$

(because $\cos(\alpha - \beta) = // remark\, 2 // = \cos(\beta - \alpha)$).

As $OX = OY = 1$, then $\overrightarrow{OX} \cdot \overrightarrow{OY} = \cos(\alpha - \beta)$ **[A1]**.

Let's write the dot product in another way. Let $\vec{e}_1, \vec{e}_2$ are unit basis vectors, $\vec{e}_1$ has coordinates $(1,0)$ and $\vec{e}_2$ has coordinates $(0,1)$, then $\overrightarrow{OX} = \cos\alpha \cdot \vec{e}_1 + \sin\alpha \cdot \vec{e}_2$ and $\overrightarrow{OY} = \cos\beta \cdot \vec{e}_1 + \sin\beta \cdot \vec{e}_2$, then

$\overrightarrow{OX} \cdot \overrightarrow{OY} = (\cos\alpha \cdot \vec{e}_1 + \sin\alpha \cdot \vec{e}_2) \cdot (\cos\beta \cdot \vec{e}_1 + \sin\beta \cdot \vec{e}_2) = (\cos\alpha \cdot \cos\beta) \cdot \vec{e}_1 \cdot \vec{e}_1 + (\cos\alpha \cdot \sin\beta) \cdot \vec{e}_1 \cdot \vec{e}_2 +$
$+ (\sin\alpha \cdot \cos\beta) \cdot \vec{e}_2 \cdot \vec{e}_1 + (\sin\alpha \cdot \sin\beta) \cdot \vec{e}_2 \cdot \vec{e}_2$. As vectors $\vec{e}_1$ and $\vec{e}_2$ are perpendicular, then

$\vec{e}_1 \cdot \vec{e}_2 = 0 = \vec{e}_2 \cdot \vec{e}_1$ and $\vec{e}_1 \cdot \vec{e}_1 = 1 = \vec{e}_2 \cdot \vec{e}_2$, therefore $\overrightarrow{OX} \cdot \overrightarrow{OY} = \cos\alpha \cdot \cos\beta + \sin\alpha \cdot \sin\beta$ **[A2]**.

From **[A1]** and **[A2]** follows **[A]**.

From **[A]** immediately follows **[B]** $\cos(\alpha + \beta) = \cos\alpha \cdot \cos\beta - \sin\alpha \cdot \sin\beta$, really:

$\cos(\alpha + \beta) = \cos(\alpha - (-\beta)) = //\ from\ [A] // = \cos\alpha \cdot \cos(-\beta) + \sin\alpha \cdot \sin(-\beta) =$
$= \cos\alpha \cdot \cos\beta - \sin\alpha \cdot \sin\beta$ and **[B]** is proved.

We have deduced two very important formulas **[A]** and **[B]**. There are also two basic formulas for $\sin(\alpha - \beta)$ and $\sin(\alpha + \beta)$. We have shown earlier that $\sin\alpha = \cos(90^o - \alpha)$ (for any angle α).
Then

$\sin(\alpha + \beta) = \cos\left(90^o - (\alpha + \beta)\right) = \cos\left((90^o - \alpha) - \beta\right) = \cos(90^o - \alpha)\cos\beta + \sin(90^o - \alpha)\sin\beta =$
$= \sin\alpha\cos\beta + \cos\alpha\sin\beta$. And we got $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$ **[C]**.

The formula for $\sin(\alpha - \beta)$ can be derived from **[C]**

$\sin(\alpha - \beta) = \sin(\alpha + (-\beta)) = \sin\alpha\cos(-\beta) + \cos\alpha\sin(-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$.

Then **[D]** $\sin(\alpha - \beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$.

Almost all the other trigonometric formulas can be very quickly derived from **[A]**,**[B]**,**[C]**,**[D]**.
For example:

$$tg(\alpha + \beta) = \frac{\sin(\alpha + \beta)}{\cos(\alpha + \beta)} = \frac{\sin\alpha\cos\beta + \cos\alpha\sin\beta}{\cos\alpha\cos\beta - \sin\alpha\sin\beta} = \frac{\dfrac{\sin\alpha\cos\beta}{\cos\alpha\cos\beta} + \dfrac{\cos\alpha\sin\beta}{\cos\alpha\cos\beta}}{\dfrac{\cos\alpha\cos\beta}{\cos\alpha\cos\beta} - \dfrac{\sin\alpha\sin\beta}{\cos\alpha\cos\beta}} = \frac{tg\alpha + tg\beta}{1 - tg\alpha \cdot tg\beta}.$$

There is also an important formula $\sin 2\alpha = \sin(\alpha+\alpha) = \sin\alpha\cos\alpha + \cos\alpha\sin\alpha = 2\sin\alpha\cos\alpha$.
And two important formulas for $\cos 2\alpha = \cos(\alpha+\alpha) = \cos\alpha\cos\alpha - \sin\alpha\sin\alpha = \cos^2\alpha - \sin^2\alpha =$
$= (1-\sin^2\alpha) - \sin^2\alpha = 1 - 2\sin^2\alpha$ **[the first one]** and
$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha = \cos^2\alpha - (1-\cos^2\alpha) = 2\cos^2\alpha - 1$ **[the second one]**.

For some reason, most of students have difficulties with trigonometry, they try to remember by heart formulas like $\cos\left(\frac{\pi}{2}+\alpha\right) = -\sin\alpha$, $ctg\left(\frac{3\pi}{2}-\alpha\right) = tg\alpha$, $\cos(\pi+\alpha) = -\cos\alpha$, despite the fact that there is no need to do that, everything can be quickly derived if we apply the basic formulas for $\sin(\alpha\pm\beta)$, $\cos(\alpha\pm\beta)$.

There is also a faster way to simplify formulas like $\cos\left(\frac{3\pi}{2}+\alpha\right)$. We just need to imagine α as a small positive angle. Then we imagine a unite circle and angles α and $\frac{3\pi}{2}+\alpha$, then we draw perpendiculars to coordinate axises in order to obtain two equal right triangles, by using these right triangles, we can quickly understand the connection between cosines and sines of given angles (in our case, the connection between cosines and sines of α and $\frac{3\pi}{2}+\alpha$).

There are also several very important formulas for a sum/difference of sines/cosines.

Let α and β are any angles. Let's show how to derive the formulas for $\sin\alpha+\sin\beta$, $\sin\alpha-\sin\beta$, $\cos\alpha+\cos\beta$, $\cos\alpha-\cos\beta$ **[T]**. The main idea here is to represent each angle α and β as a sum/difference of angles $\frac{\alpha+\beta}{2}$, $\frac{\alpha-\beta}{2}$ and to use the formulas for a sine/cosine of sum/difference. So $\alpha = \left(\frac{\alpha+\beta}{2}+\frac{\alpha-\beta}{2}\right)$, $\beta = \left(\frac{\alpha+\beta}{2}-\frac{\alpha-\beta}{2}\right)$.

Then $\sin\alpha+\sin\beta = \sin\left(\frac{\alpha+\beta}{2}+\frac{\alpha-\beta}{2}\right)+\sin\left(\frac{\alpha+\beta}{2}-\frac{\alpha-\beta}{2}\right) = \sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}+$
$+\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}+\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}-\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$.

And similarly for any other sum/difference in **[T]**.

11

Determinants

Permutations

Let's consider the set of several consecutive natural numbers $\{1,2,3\}$. The one-to-one mapping $f:1\to 2,\ 2\to 3,\ 3\to 1$ of this set onto itself can be denoted as $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ - this symbol is called a **"permutation of numbers** $1,2,3$**"** . And conversely, the permutation **[1]** $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ defines one-to-one mapping $f:1\to 1,\ 2\to 3,\ 3\to 2$ of $\{1,2,3\}$ onto itself.

Let's change positions of any two columns in the permutation $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ **[2]**-this writing is also a permutation, it defines exactly the same one-to-one mapping $f:2\to 3,\ 1\to 2,\ 3\to 1$, and we say that **[1]** and **[2]** are equal permutations, because they define the same one-to-one mapping f .

Def1. $\{1,2,3\ldots\ n\}$ are consecutive natural numbers. Let $j_1, j_2\ldots\ j_n$ are the numbers $\{1,2,3\ldots\ n\}$ which are rewritten in some order (all the numbers $j_1, j_2\ldots\ j_n$ are different, $j_k \neq j_m$ for $k \neq m$ and each j_k is some natural number from the set $\{1,2,3\ldots\ n\}$). And similarly $v_1, v_2\ldots\ v_n$ are the numbers $\{1,2,3\ldots\ n\}$ which are rewritten in some order. The symbol

[S] $\begin{pmatrix} j_1 & j_2 & j_3 & \ldots & j_n \\ v_1 & v_2 & v_3 & \ldots & v_n \end{pmatrix}$ is called a "permutation of numbers $\{1,2,3\ldots\ n\}$ ", or just a "permutation".

And **[S]** is NOT a matrix. This symbol represents one-to-one mapping of $\{1,2,3\ldots\ n\}$ onto itself: $f: j_1\to v_1,\ \ j_2\to v_2 \ldots.\ j_n \to v_n$ **[3]**. We say that two permutations are equal if they represent the same one-to-one mapping f .

Obviously, if we change positions of any two columns in **[S]** (if we permute any two columns), a new permutation will be equal to an initial one (because they both define the same one-to-one mapping). **Obviously**, two permutations are equal if and only if one of these permutations can be obtained from the other by several permutations of it's columns.

In any permutation we can always permute several colums in order to turn the first row into $1,2,3\ \ldots\ n$, we will get the permutation $\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ a_1 & a_2 & a_3 & \ldots & a_n \end{pmatrix}$ - which is called the "standard form" of an (initial) permutation .

The main idea here is that any permutation can be rewritten in a more convenient way, when it's first row consists of consecutive natural numbers $1,2,3\ \ldots\ n$. And we achieve it just by permuting several columns, so we will get exactly the same permutation as earlier, but now it is written in a "better way".

Exercise1. Show that there are exactly $n! \equiv 1 \cdot 2 \cdot \ldots \cdot n$ different permutations

$$\begin{pmatrix} j_1 & j_2 & j_3 & \cdots & j_n \\ v_1 & v_2 & v_3 & \cdots & v_n \end{pmatrix}$$

(where $j_1..j_n$ and $v_1..v_n$ are both numbers $1,2..\ n$ which are written in some order)

Hint. Use that any permutation can be written in the standard form.

Sign of a permutation. For any permutation, we say that the pair of elements (a,b) in the first row is an inversion if a is on the left of b and $a>b$. In the same way, the pair of elements (c,d) in the second row is an inversion if c is on the left of d and $c>d$.

Let's consider the permutation $\sigma \equiv \begin{pmatrix} 1 & 3 & 4 & 2 \\ 3 & 1 & 2 & 4 \end{pmatrix}$. **In the first row:** $(3,2)$ is an inversion, because 3 is on the left of 2 and $3>2$. Also $(4,2)$ is an inversion. So, there are exactly two inversions in the first row.
In the second row: here $(3,1)$ and $(3,2)$ are inversions. Then there are four inversions in σ.

Def2. If the total number of inversions in σ $\big($the number of inversions in the first row + the number of inversions in the second row$\big)$ is even, then we say that σ is an even permutation and we write $\operatorname{sgn}\sigma = 1$. If the total number of inversions in σ is odd, then we say that σ is an odd permutation, and we write $\operatorname{sgn}\sigma = -1$. (The number $\operatorname{sgn}\sigma$ is called a "sign of a permutation").

Theorem1. If some permutation σ is even/odd, then any permutation δ , which is equal to σ , is also even/odd.

Proof. Let's permute some **neighbor columns** in σ, then the number of inversions k in the first (upper) row will become $k+1$ or $k-1$, the number of inversions m in the second (lower) row will become $m+1$ or $m-1$. So, there initially were $m+k$ inversions, and now we have $m+k+2$ or $m+k-2$ or $m+k$ inversions. The parity of the total number of inversions in σ stays the same. I.e., if the total number of inversions was even, then it is still even after a permutation of any neighbor columns. And similarly if the total number of inversions was odd.
Let now we have two columns in σ and there are exactly h other columns between them.
Let's take the **right** column, we will move it to the left towards the **left** column, by permuting the **right** column with it's neighbors from the left side. Soon our columns will become neighbors and we will change their positions, then we will move the other column to the right, until it becomes on the right place. Eventually two needed columns will be permuted, it is done by several permutations of columns that stood near by. The parity of the total number of inversions does not change when we permute any two neighbor columns, therefore the parity hasn't changed.
Let $\delta = \sigma$ and σ is even or odd. We can obtain δ from σ by several permutations of columns of σ.

Each permutation of columns gives a new permutation with the same parity.
Then eventually δ has the same parity as σ.

The easiest way to determine the parity of any permutation (and therefore to determine it's sign) is to rewrite this permutation in the standard form $\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ a_1 & a_2 & a_3 & \dots & a_n \end{pmatrix}$. In such form there are no inversions at all in the first row, all the inversions may be only in the second row. So, it's easier to count the total number of inversions and to determine the sign of a permutation.

Determinant. A is $n \times n$ matrix $\mathrm{A} = \begin{pmatrix} a_{11} & .. & a_{1n} \\ .. & .. & .. \\ a_{n1} & .. & a_{nn} \end{pmatrix}$ (from now on we consider only square matrixes). For any square matrix A, the determinant of A is the number $\det \mathrm{A}$, where

$$\textbf{[V1]} \quad \det \mathrm{A} \equiv \sum_{\substack{\text{all different permutations} \\ \sigma = \begin{pmatrix} 1 & .. & n \\ k_1 & .. & k_n \end{pmatrix}}} \operatorname{sgn} \sigma \cdot a_{1k_1} \cdot \ldots \cdot a_{nk_n} \equiv \begin{bmatrix} \textit{can be} \\ \textit{denoted as} \end{bmatrix} \equiv \begin{vmatrix} a_{11} & .. & a_{1n} \\ .. & .. & .. \\ a_{n1} & .. & a_{nn} \end{vmatrix}$$

So, the determinant of $n \times n$ matrix A is a sum of $n!$ products. Each product consists of matrix elements which stay at different rows and columns, and there is a sign $+$ or $-$ in front of each product. As any permutation $\sigma = \begin{pmatrix} j_2 & j_1 & j_3 & \dots & j_n \\ v_2 & v_1 & v_3 & \dots & v_n \end{pmatrix}$ can be written in the standard form $\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ a_1 & a_2 & a_3 & \dots & a_n \end{pmatrix}$, and different permutations have different standard forms, we can rewrite the definition of $\det \mathrm{A}$ in the next way:

$$\textbf{[V2]} \quad \det \mathrm{A} = \sum_{\substack{\text{all different permutations} \\ \sigma = \begin{pmatrix} j_2 & j_1 & j_3 & \dots & j_n \\ v_2 & v_1 & v_3 & \dots & v_n \end{pmatrix}}} \operatorname{sgn} \sigma \cdot a_{j_1k_1} \cdot \ldots \cdot a_{j_nk_n}$$

In practice we never calculate concrete determinants through these definition-formulas, it's enough to remember the formulas for determinants of "small" 2×2 and 3×3 matrixes (even for 2×2 matrixes is enough). Any determinant of a bigger matrix (like 4×4 or 5×5) can be simplified and represented as a sum of determinants of smaller matrixes. There will be the examples soon.

Exercise2. Prove the next simple properties of $\det \mathrm{A}$ (by using **[V1]** or **[V2]**, which one is more convenient). **[1]** For any matrix A: $\det \mathrm{A} = \det \mathrm{A}^{\mathrm{T}}$. **[2]** If A has a row of zeroes, or a column of

zeroes, then $\det A = 0$. **[3]** If we permute any two rows in A, or any two columns in A, then a new matrix has the determinant $-\det A$.

From **[3]** follows **[3.1]**: If A has two equal columns, or two equal rows, then $\det A = 0$.

[4] If we multiply by λ all the elements of some row/column of A, the determinant of a new matrix is $\lambda \det A$.

[5] There is some determinant, where every element of some row is represented as a sum of two elements, then such determinant can be represented as a sum of two other determinants:

$$\begin{vmatrix} a_{11} & .. & .. & a_{1n} \\ b_1 + c_1 & .. & .. & b_n + c_n \\ .. & .. & .. & .. \\ a_{n1} & .. & .. & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & .. & .. & a_{1n} \\ b_1 & .. & .. & b_n \\ .. & .. & .. & .. \\ a_{n1} & .. & .. & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & .. & .. & a_{1n} \\ c_1 & .. & .. & c_n \\ .. & .. & .. & .. \\ a_{n1} & .. & .. & a_{nn} \end{vmatrix}.$$

The same representation is true for any column.

From **[5]** and **[3.1]** follows: we can take any row and add to it any linear combination of **any other** rows, a determinant will not change. And the same is true for any column.

Def. A is a **block matrix** if it looks like $A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$, here B and C are some square matrixes (and therefore A is a square matrix). Every zero 0 denotes a rectangle of zeroes.

Exercise3. Show that for any block matrix A the next is true: $\det A = \det B \cdot \det C$.

Def. A is a square matrix. Let's fix any a_{ij} and replace it by 1, in the same time all the other elements in the row with number i and in the column with number j we replace by zeroes (and all the other elements of A stay the same). The determinant of the new matrix is called **an algebraic complement of** a_{ij}, and we will denote it like A_{ij}. (**Note**: A_{ij} is NOT a matrix, it is a number. a determinant of a concrete matrix).

For example,

$\begin{vmatrix} 0 & a_{12} & .. & a_{1n} \\ 1 & 0 & .. & 0 \\ .. & .. & .. & .. \\ 0 & a_{n2} & .. & a_{nn} \end{vmatrix}$ -an algebraic complement of a_{21} and $\begin{vmatrix} 1 & 0 & .. & 0 \\ 0 & a_{22} & .. & a_{2n} \\ .. & .. & .. & .. \\ 0 & a_{n2} & .. & a_{nn} \end{vmatrix}$ -an algebraic complement of a_{11}.

Let's derive a very important formula:

$$\begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ a_{21} & a_{22} & .. & a_{2n} \\ .. & .. & .. & .. \\ a_{n1} & .. & .. & a_{nn} \end{vmatrix} \Rightarrow \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ (a_{21}+0...+0) & (0+a_{22}+0+...+0) & .. & (0+...+0+a_{2n}) \\ .. & .. & .. & .. \\ 0 & a_{n2} & .. & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ a_{21} & 0 & .. & 0 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ 0 & a_{22} & .. & 0 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix} + ... + \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ 0 & 0 & .. & a_{2n} \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix} =$$

$$= a_{21} \cdot \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ 1 & 0 & .. & 0 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix} + a_{22} \cdot \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ 0 & 1 & .. & 0 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix} + ... + a_{2n} \cdot \begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ 0 & 0 & .. & 1 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & a_{nn} \end{vmatrix}.$$

Let's consider the first determinant, we can annul in it almost all the first column. Really, let's take the first row and add to it the second row, multiplied by $-a_{11}$ (the first determinant will not change, it follows from **[5]** and **[3.1]**). Then we take the third row and we add to it the second row, multiplied by $-a_{31}$, then we take the fourth row and we add to it the second row, multiplied by $-a_{41}$ and etc.

Finally, the first determinant will look like: $\begin{vmatrix} 0 & a_{12} & .. & a_{1n} \\ 1 & 0 & .. & 0 \\ .. & .. & .. & .. \\ 0 & a_{n2} & .. & a_{nn} \end{vmatrix} = \mathrm{A}_{21}$ - it is an algebraic complement of a_{21}. And similarly, for any other determinant in our sum, then:

$$\begin{vmatrix} a_{11} & a_{12} & .. & a_{1n} \\ a_{21} & a_{22} & .. & a_{2n} \\ .. & .. & .. & .. \\ a_{n1} & .. & .. & a_{nn} \end{vmatrix} = a_{21} \cdot \begin{vmatrix} 0 & a_{12} & .. & a_{1n} \\ 1 & 0 & .. & 0 \\ .. & .. & .. & .. \\ 0 & a_{n2} & .. & a_{nn} \end{vmatrix} + a_{22} \cdot \begin{vmatrix} a_{11} & 0 & .. & a_{1n} \\ 0 & 1 & .. & 0 \\ .. & .. & .. & .. \\ a_{n1} & 0 & .. & a_{nn} \end{vmatrix} + ... + a_{2n} \cdot \begin{vmatrix} a_{11} & a_{12} & .. & 0 \\ 0 & 0 & .. & 1 \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & .. & 0 \end{vmatrix} \Leftrightarrow$$

$\Leftrightarrow \det \mathrm{A} = a_{21} \cdot \mathrm{A}_{21} + ... + a_{2n} \cdot \mathrm{A}_{2n}$. The same formula is obviously true for any other row (with number i), i.e., $\det \mathrm{A} = a_{i1} \cdot \mathrm{A}_{i1} + ... + a_{in} \cdot \mathrm{A}_{in}$ **[D]**.

And similarly for any column (with number j) $\det \mathrm{A} = a_{1j} \cdot \mathrm{A}_{1j} + ... + a_{nj} \cdot \mathrm{A}_{nj}$ **[E]**.

These formulas are not our final formulas, so we need to continue.

Def. Let's fix any element a_{ij} of A and delete the row number i and the column number j.

We will get $(n-1)\times(n-1)$ matrix Λ_{ij}, the determinant of that matrix must be denoted as

$\Delta_{ij} = \det \Lambda_{ij}$

Assertion1. $\mathrm{A}_{ij} = (-1)^{i+j}\cdot\Delta_{ij}$ for any i, j.

Proof. Let's take some determinant A_{ij}, we take it's column with number j (it consists of several zeroes and 1) we will consecutively permute this column with it's left neighbor columns.

Finally this column will become the first one, and the row with number i will look like $(1,0,0\ldots0)$. We will permute this row with it's upper neighbor rows. Finally this row will become the first one

and the determinant will look like: $\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \ldots & & \Lambda_{ij} & \\ 0 & & & \end{vmatrix}$. This determinant is obtained from A_{ij} by $j-1$ consecutive permutations of columns and $i-1$ consecutive permutations of rows.

Each permutation of any two rows/columns changes the sign of a determinant, therefore

$\mathrm{A}_{ij}\cdot(-1)^{j-1+i-1} = \begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \ldots & & \Lambda_{ij} & \\ 0 & & & \end{vmatrix}$ - there is a block matrix on the right side.

According to the **exercise3**, the right determinant is equal to $\det\Lambda_{ij} = \Delta_{ij}$. Therefore

$\mathrm{A}_{ij}\cdot(-1)^{j-1+i-1} = \Delta_{ij} \Rightarrow \mathrm{A}_{ij}\cdot(-1)^{j+i} = \Delta_{ij} \Rightarrow \mathrm{A}_{ij} = (-1)^{j+i}\cdot\Delta_{ij}$.

Then from **[D]** we have: $\det \mathrm{A} = a_{i1}\cdot(-1)^{i+1}\Delta_{i1} + \ldots + a_{in}\cdot(-1)^{i+n}\Delta_{in}$ **[D1].**

And similarly, from **[E]** we have: $\det \mathrm{A} = a_{1j}\cdot(-1)^{1+j}\Delta_{1j} + \ldots + a_{nj}\cdot(-1)^{n+j}\Delta_{nj}$ **[E1].**

Notice that any Δ_{ij} is a determinant of a matrix which is smaller than A. Formulas **[D1]** and **[E1]** are our basic formulas for determinants-calculation.

The simpliest determinant: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Let's calculate $\begin{vmatrix} a & b & c \\ d & e & f \\ g & l & p \end{vmatrix}$ by using **[D1]** (we take the first row).

Then

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & l & p \end{vmatrix} = a\cdot(-1)^{1+1}\cdot\begin{vmatrix} e & f \\ l & p \end{vmatrix} + b\cdot(-1)^{1+2}\cdot\begin{vmatrix} d & f \\ g & p \end{vmatrix} + c\cdot(-1)^{1+3}\cdot\begin{vmatrix} d & e \\ g & l \end{vmatrix} = a(ep-fl)-b(dp-fg)+c(dl-eg)=$$

$= aep + bfg + cdl - ceg - bdp - afl$. It's very easy to remember this formula if we notice the geometrical pattern (the positions of our factors in the initial 3×3 matrix).

So $\begin{vmatrix} a & b & c \\ d & e & f \\ g & l & p \end{vmatrix} = aep + bfg + cdl - ceg - bdp - afl$.

Example1. Examples of determinants:

(a) $\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = 1\cdot 3 - 2\cdot 2 = 3 - 4 = -1$, (b) $\begin{vmatrix} -2 & 4 \\ 3 & 4 \end{vmatrix} = (-2)\cdot 4 - 4\cdot 3 = -8 - 12 = -20$,

(c) $\begin{vmatrix} -1 & 3 \\ -2 & 3 \end{vmatrix} = (-1)\cdot 3 - 3\cdot(-2) = -3 + 6 = 3$, (d) $\begin{vmatrix} 1 & 0 & -2 \\ 2 & 2 & 3 \\ 1 & -1 & 4 \end{vmatrix}$ the second column contains zero,

so let's expand this determinant along the second column:

$$\begin{vmatrix} 1 & 0 & -2 \\ 2 & 2 & 3 \\ 1 & -1 & 4 \end{vmatrix} = 0\cdot(-1)^{1+2}\cdot\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} + 2\cdot(-1)^{2+2}\cdot\begin{vmatrix} 1 & -2 \\ 1 & 4 \end{vmatrix} + (-1)\cdot(-1)^{3+2}\cdot\begin{vmatrix} 1 & -2 \\ 2 & 3 \end{vmatrix} =$$

$= 2\cdot(1\cdot 4 - (-2)\cdot 1) + 1\cdot(1\cdot 3 - (-2)\cdot 2) = 12 + 7 = 19.$

(e) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ let's expand along the first row:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1(-1)^{1+1}\cdot\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} + 2(-1)^{1+2}\cdot\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3\cdot(-1)^{1+3}\cdot\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} =$$

$= 1\cdot(5\cdot 9 - 6\cdot 8) - 2\cdot(4\cdot 9 - 6\cdot 7) + 3\cdot(4\cdot 8 - 5\cdot 7) = -3 + 12 - 9 = 0.$

We could also understand that the determinant (e) is zero almost without any calculations. In any matrix we can take any row/column and add to it any linear combination of other rows/columns, a determinant will not change. In particular, we can take any row/column and subtract from it any other row/column. We have $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$, let's take the second row and subtract the first row from it:

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}=\begin{vmatrix}1&2&3\\4-1&5-2&6-3\\7&8&9\end{vmatrix}=\begin{vmatrix}1&2&3\\3&3&3\\7&8&9\end{vmatrix}=\begin{vmatrix}1&2&3\\3&3&3\\7-1&8-2&9-3\end{vmatrix}=\begin{vmatrix}1&2&3\\3&3&3\\6&6&6\end{vmatrix}=2\cdot\begin{vmatrix}1&2&3\\3&3&3\\3&3&3\end{vmatrix}=2\cdot 0=0$$

Example2. (f) $\begin{vmatrix}1&-1&3&4\\2&2&2&4\\3&3&4&6\\4&9&8&10\end{vmatrix}$ - it is a determinant 4×4. It's always very important to notice if we can make some elementary manipulations with rows and columns to simplify the determinant (like to take some row/column and add/subtract to/from it some other row/column), we want to make so many zeroes as possible in some row/column, and after we will use **[D1]/[E1]**.

$$\begin{vmatrix}1&-1&3&4\\2&2&2&4\\3&3&4&7\\4&9&8&10\end{vmatrix}=\begin{vmatrix}1&-1&(3+1)&4\\2&2&(2+2)&4\\3&3&(4+3)&7\\4&9&(8+4)&10\end{vmatrix}=\begin{vmatrix}1&-1&4&4\\2&2&4&4\\3&3&7&7\\4&9&12&10\end{vmatrix}=\begin{vmatrix}1&-1&4&(4-4)\\2&2&4&(4-4)\\3&3&7&(7-7)\\4&9&12&(10-12)\end{vmatrix}=\begin{vmatrix}1&-1&4&0\\2&2&4&0\\3&3&7&0\\4&9&12&-2\end{vmatrix}$$

now we can expand this determinant along the last column

$$\begin{vmatrix}1&-1&4&0\\2&2&4&0\\3&3&7&0\\4&9&12&-2\end{vmatrix}=-2\cdot(-1)^{4+4}\cdot\begin{vmatrix}1&-1&4\\2&2&4\\3&3&7\end{vmatrix}=-2\cdot\begin{vmatrix}1&-1&4\\2&2&4\\3&3&7\end{vmatrix}=$$

$$=-2\cdot\begin{vmatrix}1&-1&4\\2+2\cdot 1&2+2\cdot(-1)&4+2\cdot 4\\3&3&7\end{vmatrix}=-2\cdot\begin{vmatrix}1&-1&4\\4&0&12\\3&3&7\end{vmatrix}$$ and finally, we expand this determinant along the second column

$$-2\cdot\begin{vmatrix}1&-1&4\\4&0&12\\3&3&7\end{vmatrix}=-2\cdot\left((-1)\cdot(-1)^{1+2}\cdot\begin{vmatrix}4&12\\3&7\end{vmatrix}+3\cdot(-1)^{3+2}\cdot\begin{vmatrix}1&4\\4&12\end{vmatrix}\right)=-2\big(1\cdot(4\cdot 7-12\cdot 3)-3(12-16)\big)=$$

$$-2(-8+12)=-8.$$

Def: a matrix A is called diagonal if $a_{ij}=0$ for any $i\neq j$. The only non-zero elements in such matrix may stay on it's main diagonal $a_{11}, a_{22} \dots a_{nn}$.

$$A=diag\{a_{11},a_{22}\dots a_{nn}\}\Leftrightarrow A=\begin{pmatrix}a_{11}&..&0\\..&...&..\\0&...&a_{nn}\end{pmatrix}.$$

The matrix $E=diag\{1,1\dots 1\}$ is called a "unit matrix".

If all the elements below the main diagonal are zeroes ($a_{ij}=0$ for any $i>j$), then A is called upper triangular.

If all the elements under the main diagonal are zeroes ($a_{ij} = 0$ for any $i < j$), then A is called lower triangular.
If a matrix is lower triangular **or** upper triangular, then we call it just “triangular”.

Exercise4. Show (by using **[V1]** or **[D1]** or **[E1]**) that the determinant of any triangular matrix is a product of it’s diagonal elements. So, A is triangular $\Rightarrow \det \mathrm{A} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$.

In particular $\det \mathrm{E} = 1$.

The main theorem. For any $n \times n$ matrixes A and B: $\det(\mathrm{A} \cdot \mathrm{B}) = \det \mathrm{A} \cdot \det \mathrm{B}$.

This theorem is very important, and we will provide a very simple proof for it. But at first we need to introduce some simple theory, which will allow us to change our approach to determinants.

Linear spaces

Def. F is a field and L is a commutative additive group. There exist the operation ("multiplication") $"\cdot": F \times L \to L$ that for every pair $\lambda, \vec{v} \,\|\, \lambda \in F, \vec{v} \in L$ compares some element $\lambda \cdot \vec{v} \in L$. And $"\cdot"$ has the next properties:

[A] $1 \cdot \vec{v} = \vec{v} \,\|\, \forall \vec{v} \in L$ and $\lambda \cdot (\mu \cdot \vec{v}) = (\lambda \cdot \mu) \cdot \vec{v} \,\|\, \forall \lambda, \mu \in F, \forall \vec{v} \in L$.

[B] "sort of distributivity" $(\lambda + \mu) \cdot \vec{v} = \lambda \cdot \vec{v} + \mu \cdot \vec{v}$ and $\lambda \cdot (\vec{v} + \vec{m}) = \lambda \cdot \vec{v} + \lambda \cdot \vec{m}$ for any $\lambda, \mu \in F, \forall \vec{v}, \vec{m} \in L$.

Then L is called a linear space under the field F. Elements of L are called "vectors". And elements of F are called "scalars". We can just say "L is a space under F" or "L is a space".

The most important example is the linear space of $3-d$ vectors which was built earlier.
Notice, in mathematical model equal vectors must be considered as the same vector.
So when we say "a vector" we imply the whole class of equal vectors. $3-d$ vectors form a commutative group L, we can multiply vectors by real numbers R, so we have a linear space under R, it's easy to see that **[A]** and **[B]** are true.

Let's notice that the set of all $m \times n$ matrixes which consist of real numbers, form a vector space under R (in this case every matrix is a "vector" and any real number is a "scalar").

Two most important vector spaces (of matrixes) are: the linear space R^n of $n \times 1$ columns $\left\{ \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} \right\}$

and the linear space R_n of $1 \times n$ rows $\{(a_1 \quad \dots \quad a_n)\}$.

Def. L is a linear space under F. For any vectors $\vec{v}_1, \vec{v}_2 \dots \vec{v}_n \in L$ and any scalars $\lambda_1, \lambda_2 \dots \lambda_n \in F$ the vector $\lambda_1 \cdot \vec{v}_1 + \lambda_2 \cdot \vec{v}_2 + \dots + \lambda_n \cdot \vec{v}_n$ is called a "linear combination of $\vec{v}_1, \vec{v}_2 \dots . \vec{v}_n$ with coefficients $\lambda_1, \lambda_2 \dots . \lambda_n$" or just a "linear combination of $\vec{v}_1, \vec{v}_2 \dots . \vec{v}_n$".

Def. L is a linear space. If there exist some set of vectors $\{\vec{e}_1 \dots \vec{e}_n\} \subset L$ such that any vector $\vec{v} \in L$ can be represented as a linear combination of these vectors $\vec{v} = \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n$ and for any vector $\vec{v}$ it's representation is **unique**, then the set $\{\vec{e}_1 \dots \vec{e}_n\}$ is called a basis of the space L.

Example. Prove that the **"standard basis vectors"** $\vec{e}_1 = (1,0 \dots 0)$ $\vec{e}_2 = (0,1 \dots 0) \dots . \vec{e}_n = (0,0 \dots 1)$ form a basis of R_n and the "standard basis vectors" $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ .. \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ .. \\ 0 \end{pmatrix} \dots . \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ .. \\ 1 \end{pmatrix}$ form a basis of R^n .

Exercise. Provide some basis of the space of $m \times n$ matrixes.

From now on, L is always a linear space with some fixed basis $\{\vec{e}_1 \ldots \vec{e}_n\} \subset L$.

And L is a linear space under the field R of real numbers.

Def. L is a linear space. Every rule f that for any vector $\vec{v} \in L$ compares some real number $f(\vec{v}) \in R$ is called a form on L.

For any vector $\vec{v}$, the number $f(\vec{v})$ is called "the value of f on $\vec{v}$".

Two forms f, g, which are defined on L, are equal if $f(\vec{v}) = g(\vec{v})$ for any $\vec{v} \in L$. The form f is called linear if **[A]** $f(\lambda\vec{v}) = \lambda f(\vec{v})$ for any $\vec{v} \in L$ and any $\lambda \in R$

[B] $f(\vec{v} + \vec{u}) = f(\vec{v}) + f(\vec{u})$ for any $\vec{v}, \vec{u} \in L$.

Assertion1. f is a linear form. Then the value of f on any vector $\vec{v}$ is uniquely defined by the values of this form on the basis vectors $\vec{e}_1, \vec{e}_2 \ldots \vec{e}_n$.

Proof. Really, for any vector $\vec{v} \in L$ there is a unique representation $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \ldots + v_n\vec{e}_n$, then $f(\vec{v}) = v_1 f(\vec{e}_1) + v_2 f(\vec{e}_2) + \ldots + v_n f(\vec{e}_n)$, then the assertion1 is true.

Therefore: two linear forms f, g are equal $\Leftrightarrow$ they are equal on the basis vectors $\vec{e}_1, \vec{e}_2 \ldots \vec{e}_n$: $f(\vec{e}_k) = g(\vec{e}_k)$ for any k.

Def. Every rule f that for any ordered set of vectors $(\vec{v}_1, \vec{v}_2 \ldots \vec{v}_n)$ (every vector $\vec{v}_k$ is taken from L) compares some real number $f(\vec{v}_1, \vec{v}_2 \ldots \vec{v}_n) \in R$ is called an "n-form on L" or just "a form on L".

A form f is called **normalized** if $f(\vec{e}_1, \vec{e}_2 \ldots \vec{e}_n) = 1$.

f is called **polylinear** if f is linear in all it's arguments, i.e., for any number $T \in [1, n]$:
$f(\vec{v}_1 .. \vec{v}_T + \vec{u}_T .. \vec{v}_n) = f(\vec{v}_1 .. \vec{v}_T .. \vec{v}_n) + f(\vec{v}_1 .. \vec{u}_T .. \vec{v}_n)$ and $f(\vec{v}_1 .. \lambda\vec{v}_T .. \vec{v}_n) = \lambda \cdot f(\vec{v}_1 .. \vec{v}_T .. \vec{v}_n)$.

f is called **symmetric** if it doesn't change when we permute any two of it's arguments.

For any $T, J \in [1, n] : T \neq J$ there must be $f(\vec{v}_1 .. \vec{v}_T .. \vec{v}_J .. \vec{v}_n) = f(\vec{v}_1, .. \vec{v}_J .. \vec{v}_T .. \vec{v}_n)$.

And f is called **skew-symmetric** if it changes the sign when we permute any two of it's arguments.

For any $T, J \in [1, n] : T \neq J$ there must be $f(\vec{v}_1 .. \vec{v}_T .. \vec{v}_J .. \vec{v}_n) = -f(\vec{v}_1, .. \vec{v}_J .. \vec{v}_T .. \vec{v}_n)$.

Let A is $n \times n$ matrix. Rows of A are some vectors from R_n.

The determinant $\det \mathrm{A}$ of any matrix A is a real number. Any matrix A can be considered as exactly n vectors from R_n, which are the $1-st,\ 2-nd,\ ..\ n-th$ rows of A.

Then $\det \mathrm{A}$ is an n-form on the space R_n, really

$\det A = f(\vec{v}_1 .. \vec{v}_n) = \begin{vmatrix} components\ of & \vec{v}_1 \\ .... & ...... \\ components\ of & \vec{v}_n \end{vmatrix}$. For any vectors $\vec{v}_1, \vec{v}_2 .. \vec{v}_n$ we have: $f(\vec{v}_1 .. \vec{v}_n) \equiv \det A$ is a concrete real number. Then $\det A$ is really an n-form on R_n.

According to the simpliest properties of determinant, the form $f \equiv \det A$ (on R_n) is **normalized** (because $\det E = 1$), polylinear and skew-symmetric.

From now on, $(k_1, k_2 ... k_n)$ are the numbers $1, 2 ... n$ which are rewritten in some order.

Let $\operatorname{sgn}(k_1, k_2 ... k_n) \equiv 1$ if the total amount of inversions in the row $(k_1, k_2 ... k_n)$ is even and $\operatorname{sgn}(k_1, k_2 ... k_n) \equiv -1$ if the total amount of inversions in the row $(k_1, k_2 ... k_n)$ is odd.

Let's notice that $\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 & ... & n \\ k_1 & k_2 & k_3 & ... & k_n \end{pmatrix} = \operatorname{sgn}(k_1, k_2 ... k_n)$, because there are no inversions at all in the first row.

Theorem2. There exist exactly one n -form on R_n which is normative, polylinear and skew-symmetric.

Proof. The existence is proved (the determinant is such form). Let's prove the uniqueness.

Let $f(\vec{v}_1 .. \vec{v}_n)$ is some normative, polylinear and skew-symmetric form on R_n. As f is polylinear, it changes the sign when we permute any two of it's arguments. It's easy to understand that from here follows: $f(\vec{v}_{k_1} .. \vec{v}_{k_n}) = \operatorname{sgn}(k_1 .. k_n) \cdot f(\vec{v}_1 .. \vec{v}_n)$ **[E]**. Let's notice, as f is normative, for the basis vectors $\vec{e}_1 .. \vec{e}_n$ the equality **[E]** looks like

$f(\vec{e}_{k_1} .. \vec{e}_{k_n}) = \operatorname{sgn}(k_1 .. k_n) \cdot f(\vec{e}_1 .. \vec{e}_n) \Rightarrow f(\vec{e}_{k_1} .. \vec{e}_{k_n}) = \operatorname{sgn}(k_1 .. k_n)$ **[E1]**.

Each vector $\vec{v}_1 .. \vec{v}_n$ can be uniquely represented as a linear combination of basis vectors:

$$\vec{v}_1 = a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + ... + a_{1n}\vec{e}_n \,\|\, \vec{v}_2 = a_{21}\vec{e}_1 + a_{22}\vec{e}_2 + ... + a_{2n}\vec{e}_n \,\| .... \|\, \vec{v}_n = a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + ... + a_{nn}\vec{e}_n \,\|$$

then $f(\vec{v}_1 .. \vec{v}_n) = f(a_{11}\vec{e}_1 + a_{12}\vec{e}_2 + ... + a_{1n}\vec{e}_n \ ,..., \ a_{n1}\vec{e}_1 + a_{n2}\vec{e}_2 + ... + a_{nn}\vec{e}_n) = [linearity] =$

$$= \sum_{\substack{all\ different\ sets \\ (k_1, k_2 ... k_n)}} a_{1k_1} \cdot a_{2k_2} \cdot ... \cdot a_{nk_n} \cdot f(\vec{e}_{k_1} .. \vec{e}_{k_n}) = //[E1]// =$$

$$= \sum_{\substack{all\ different\ sets \\ (k_1, k_2 ... k_n)}} a_{1k_1} \cdot a_{2k_2} \cdot ... \cdot a_{nk_n} \cdot \operatorname{sgn}(k_1, k_2 .. k_n) = \sum_{\substack{all\ different\ permutations \\ \sigma = \begin{pmatrix} 1 & .. & n \\ k_1 & .. & k_n \end{pmatrix}}} a_{1k_1} \cdot ... \cdot a_{nk_n} \cdot \operatorname{sgn} \sigma$$

we came to exactly the same formula that we had in the determinant-definition.

So, the value of our form $f(\vec{v}_1 .. \vec{v}_n)$ on any set of vectors $\vec{v}_1 .. \vec{v}_n \in R_n$ is exactly the same as a determinant of the matrix which is formed from these vectors $\begin{vmatrix} - & vector & \vec{v}_1 & - \\ - & vector & \vec{v}_2 & - \\ - & \ldots & \ldots & - \\ - & vector & \vec{v}_n & - \end{vmatrix}$.

The uniqueness is proved.

Theorem3. f and g are both **polylinear** and **skew-symmetric** forms on R_n.
If $f(\vec{e}_1 .. \vec{e}_n) = g(\vec{e}_1 .. \vec{e}_n)$, then $f \equiv g$ on R_n.

Proof. We have derived above the formula for any polylinear form:

$$f(\vec{v}_1 .. \vec{v}_n) = \sum_{\substack{all\ sets \\ (k_1, k_2 .. k_n)}} a_{1k_1} \cdot a_{2k_2} \cdot \ldots \cdot a_{nk_n} \cdot f(\vec{e}_{k_1} .. \vec{e}_{k_n}) = [skew-symmetricity] =$$

$$= \sum_{\substack{all\ sets \\ (k_1, k_2 ... k_n)}} a_{1k_1} \cdot a_{2k_2} \cdot \ldots \cdot a_{nk_n} \cdot f(\vec{e}_1 .. \vec{e}_n) \cdot \operatorname{sgn}(k_1, k_2 \ldots k_n),$$ from here we see that the value of the form f on any set of vectors $(\vec{v}_1 .. \vec{v}_n)$ is uniquely defined by the number $f(\vec{e}_1 .. \vec{e}_n)$.
And exactly the same formula is true for g. So if $f(\vec{e}_1 .. \vec{e}_n) = g(\vec{e}_1 .. \vec{e}_n)$, then $f \equiv g$ on R_n.

Let's prove now the **main theorem**: $\det(A \cdot B) = \det A \cdot \det B$.

Let $\beta_1 \ldots \beta_n$ are the columns of B, then $A \cdot B$ has the columns: $A \cdot \beta_1 \ \ldots \ A \cdot \beta_n$ (it follows from the definition of matrix multiplication). So, we can write: $B = (\beta_1 \ldots \beta_n)$ and $A \cdot B = (A \cdot \beta_1 \ \ldots \ A \cdot \beta_n)$.

We want to prove that: $\det(A \cdot B) = \det A \cdot \det B \Leftrightarrow \det(A \cdot \beta_1 \ \ldots \ A \cdot \beta_n) = \det A \cdot \det(\beta_1 \ldots \beta_n)$.

Let $\vec{v}_1 .. \vec{v}_n$ are any vector-columns from R^n. Let's denote $\det(A \cdot \vec{v}_1 \ \ldots \ A \cdot \vec{v}_n) \equiv f(\vec{v}_1 \ldots \vec{v}_n)$ - it is obviously a polylinear and skew symmetric form and vectors $\vec{v}_1 \ldots \vec{v}_n$ are it's arguments.
And $\det A \cdot \det(\vec{v}_1 \ldots \vec{v}_n) \equiv g(\vec{v}_1 \ldots \vec{v}_n)$ is also a polylinear and skew symmetric form with arguments $\vec{v}_1 \ldots \vec{v}_n$. It's easy to see that $f(\vec{e}_1 \ldots \vec{e}_n) = g(\vec{e}_1 \ldots \vec{e}_n) = \det A$ then, according to the **theorem3**, $f \equiv g$ on any set of vectors $\vec{v}_1 \ldots \vec{v}_n$ from R^n, in particular for $\beta_1 \ldots \beta_n$ we have:

$$f(\beta_1 \ldots \beta_n) = g(\beta_1 \ldots \beta_n) \Leftrightarrow \det(A \cdot B) = \det A \cdot \det B.$$

Inverse matrix

Def: a matrix A is called **singular** if $\det A = 0$ and **nonsingular** if $\det A \neq 0$.

Def. A is $n \times n$ matrix. If there exist some $n \times n$ matrix B such that $A \cdot B = B \cdot A = E$, then B is called an inverse matrix, and we denote $B \equiv A^{-1}$. So $A \cdot A^{-1} = A^{-1} \cdot A = E$.

Comment1. If A is singular, then an inverse matrix A^{-1} does not exist.

Proof. Let's assume that A^{-1} exists, then $A^{-1} \cdot A = E$, then (the main theorem) $\det A^{-1} \cdot \det A = \det E \Rightarrow \det A^{-1} \cdot 0 = 1 \Leftrightarrow 0 = 1$ and we have a contradiction. Then A^{-1} doesn't exist.

Def. Let A is $m \times n$ matrix. The transposed matrix A^T is $n \times m$ matrix C such that $c_{ij} = a_{ji} \forall i, j$

Put simply: rows of A^T are the columns of A and columns of A^T are the rows of A.

Examples.

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \parallel A = \begin{pmatrix} 1 & 2 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \parallel A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \parallel$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \parallel A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \parallel A = \begin{pmatrix} 2 & 3 & 5 \\ 0 & 1 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 2 & 0 \\ 3 & 1 \\ 5 & 6 \end{pmatrix}.$$

When we go from A to A^T we say that we "transpose our matrix".

Exercise. Prove the next basic properties: **[1]** $(A^T)^T = A$ for any matrix A,

[2] $(A \cdot B)^T = B^T \cdot A^T$ for any $m \times n$ and $n \times k$ matrixes A and B.

Theorem4. For any nonsingular matrix A there exist a unique inverse matrix A^{-1}.

Let's prove at first the existence of A^{-1}.

Lemma1. If we multiply any row of A by algebraic complements of any other row, we will get zero.

Comment. Let's take for example the first row $a_{11}, a_{12} \dots a_{1n}$ and $A_{21}, A_{22} \dots A_{2n}$ are algebraic complements of the second row, then there must be $a_{11}A_{21} + a_{12}A_{22} + \dots + a_{1n}A_{2n} = 0$.

Let's take the fourth row $a_{41}, a_{42} \dots a_{4n}$ and $A_{11}, A_{12} \dots A_{1n}$ are algebraic complements of the first row, then $a_{41}A_{11} + a_{42}A_{12} + \dots + a_{4n}A_{1n} = 0$ and etc.

Let's show for example that $a_{11}A_{21} + a_{12}A_{22} + \dots + a_{1n}A_{2n} = 0$.

Let's take the determinant $\det A$ and replace it's second row by it's first row:

$$\begin{vmatrix} a_{11} & a_{12} & .. & .. & a_{1n} \\ a_{11} & a_{12} & .. & .. & a_{1n} \\ a_{31} & a_{32} & .. & .. & a_{3n} \\ .. & ... & .. & .. & .. \\ a_{n1} & a_{n2} & .. & .. & a_{nn} \end{vmatrix}$$ **[T]** then the sum $a_{11}A_{21}+a_{12}A_{22}+\dots+a_{1n}A_{2n}$ is exactly the determinant **[T]**.

And **[T]** is a determinant of a matrix with two equal rows, such determinant is equal to zero, then $a_{11}A_{21}+a_{12}A_{22}+\dots+a_{1n}A_{2n}=0$. The same is true for columns: if we multiply any column of A by algebraic complements of any other column, we will get zero.

Let's build A^{-1} now.

[step1] We take the initial matrix A and instead of every element a_{ij} we write it's algebraic complement: $a_{ij} \to A_{ij} = (-1)^{i+j} \cdot \Delta_{ij}$.

[step2] Every element of the new matrix must be divided by $\det A$. I.e.: $A_{ij} \to \dfrac{A_{ij}}{\det A}$.

[step3] We transpose the matrix from the **step2**. The new matrix is exactly A^{-1}

$$\begin{vmatrix} a_{11} & a_{12} & .. & .. & a_{1n} \\ a_{11} & a_{22} & .. & .. & a_{2n} \\ a_{31} & a_{32} & .. & .. & a_{3n} \\ .. & ... & .. & .. & .. \\ a_{n1} & a_{n2} & .. & .. & a_{nn} \end{vmatrix} \xrightarrow[a_{ij}\to A_{ij}]{} \begin{vmatrix} A_{11} & A_{12} & .. & .. & A_{1n} \\ A_{21} & A_{22} & .. & .. & A_{2n} \\ A_{31} & A_{32} & .. & .. & A_{3n} \\ .. & ... & .. & .. & .. \\ A_{n1} & A_{n2} & .. & .. & A_{nn} \end{vmatrix} \xrightarrow[A_{ij}\to A_{ij}/\det A]{}$$

$$\xrightarrow[A_{ij}\to A_{ij}/\det A]{} \begin{vmatrix} A_{11}/\det A & A_{12}/\det A & .. & .. & A_{1n}/\det A \\ A_{21}/\det A & A_{22}/\det A & .. & .. & A_{2n}/\det A \\ A_{31}/\det A & A_{32}/\det A & .. & .. & A_{3n}/\det A \\ .. & ... & .. & .. & .. \\ A_{n1}/\det A & A_{n2}/\det A & .. & .. & A_{nn}/\det A \end{vmatrix} \xrightarrow[Transpose]{}$$

$$\xrightarrow[Transpose]{} \begin{vmatrix} A_{11}/\det A & A_{21}/\det A & .. & .. & A_{n1}/\det A \\ A_{12}/\det A & A_{22}/\det A & .. & .. & A_{n2}/\det A \\ A_{13}/\det A & A_{23}/\det A & .. & .. & A_{n3}/\det A \\ .. & ... & .. & .. & .. \\ A_{1n}/\det A & A_{2n}/\det A & .. & .. & A_{nn}/\det A \end{vmatrix} \equiv A^{-1}.$$

Now, by using the lemma1, we can easily check that

$$A\cdot A^{-1}=\begin{vmatrix} a_{11} & a_{12} & .. & .. & a_{1n}\\ a_{11} & a_{22} & .. & .. & a_{2n}\\ a_{31} & a_{32} & .. & .. & a_{3n}\\ .. & ... & .. & .. & ..\\ a_{n1} & a_{n2} & .. & .. & a_{nn}\end{vmatrix}\cdot\begin{vmatrix} A_{11}/\det A & A_{21}/\det A & .. & .. & A_{n1}/\det A\\ A_{12}/\det A & A_{22}/\det A & .. & .. & A_{n2}/\det A\\ A_{13}/\det A & A_{23}/\det A & .. & .. & A_{n3}/\det A\\ .. & ... & .. & .. & ..\\ A_{1n}/\det A & A_{2n}/\det A & .. & .. & A_{nn}/\det A\end{vmatrix}=E,$$

and also that

$$A^{-1}\cdot A=\begin{vmatrix} A_{11}/\det A & A_{21}/\det A & .. & .. & A_{n1}/\det A\\ A_{12}/\det A & A_{22}/\det A & .. & .. & A_{n2}/\det A\\ A_{13}/\det A & A_{23}/\det A & .. & .. & A_{n3}/\det A\\ .. & ... & .. & .. & ..\\ A_{1n}/\det A & A_{2n}/\det A & .. & .. & A_{nn}/\det A\end{vmatrix}\cdot\begin{vmatrix} a_{11} & a_{12} & .. & .. & a_{1n}\\ a_{11} & a_{22} & .. & .. & a_{2n}\\ a_{31} & a_{32} & .. & .. & a_{3n}\\ .. & ... & .. & .. & ..\\ a_{n1} & a_{n2} & .. & .. & a_{nn}\end{vmatrix}=E.$$

The existence is proved.

Uniqueness. Let's assume that there exist some other inverse matrix B such that $A\cdot B=B\cdot A=E$. We already have $A\cdot A^{-1}=A^{-1}\cdot A=E$. Let's multiply both sides of $A\cdot B=E$ by A^{-1} from the left side, then

$A^{-1}\cdot(A\cdot B)=A^{-1}\cdot E \Leftrightarrow A^{-1}\cdot(A\cdot B)=A^{-1}\Leftrightarrow[\textit{associativ ity of matrix multiplication}]\Leftrightarrow$

$\Leftrightarrow(A^{-1}\cdot A)\cdot B=A^{-1}\Rightarrow B=A^{-1}$. So, for any nonsingular matrix A an inverse matrix A^{-1} is unique.

Properties of an inverse matrix. [A] $\det A^{-1}=1/\det A$. Really $A\cdot A^{-1}=A^{-1}\cdot A=E$, then $A^{-1}\cdot A=E\Rightarrow\det(A^{-1})\cdot\det(A)=\det E\Rightarrow\det(A^{-1})\cdot\det(A)=1$.

[B] If $A\cdot B=E$ or $B\cdot A=E$, then $B=A^{-1}$ (the proof is the same as in the uniqueness).

[C] A is a nonsingular matrix. If for some matrixes X,Y we have $AX=AY$, then $X=Y$

[D] For any nonsingular matrixes A,B: $(A\cdot B)^{-1}=B^{-1}\cdot A^{-1}$.

Proof. [C] is obvious. Let's prove **[D]**. By definition of an inverse matrix $(A\cdot B)\cdot(A\cdot B)^{-1}=/\textit{by def}/=E$. Let's calculate $(A\cdot B)\cdot B^{-1}\cdot A^{-1}$. As matrix multiplication is associative, we can rearrange brackets in any way we want:

$(A\cdot B)\cdot B^{-1}\cdot A^{-1}=A\cdot(B\cdot B^{-1})\cdot A^{-1}=A\cdot(E)\cdot A^{-1}=A\cdot A^{-1}=E$. Then: if we multiply a nonsingular matrix $A\cdot B$ by any of the matrixes $(A\cdot B)^{-1}$, $B^{-1}\cdot A^{-1}$ from the right side, we will get the unit matrix E. Then, according to **[C]**, there must be $(A\cdot B)^{-1}=B^{-1}\cdot A^{-1}$.

[E] For any nonsingular matrix A: $(A^{-1})^T = (A^T)^{-1}$.

Let's show that if we multiply a nonsingular matrix A^T by any of these matrixes from the right side, we will get the unit matrix E (then, by **[C]** these matrixes are equal). So,

$A^T \cdot (A^T)^{-1} = /\,\textit{by def}\,/ = E$ and

$A^T \cdot (A^{-1})^T = //\,\textit{use the equality}\,(AB)^T = B^T \cdot A^T\,// = (A^{-1} \cdot A)^T = (E)^T = E$.

Examples. Find the inverse matrix for $A = \begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix}$.

[step1] $a_{11} = 1 \to A_{11} = (-1)^{1+1} \cdot 15 = 15 \,\|\, a_{12} = 8 \to A_{12} = (-1)^{1+2} \cdot 2 = -2 \,\|\, a_{21} = 1 \to A_{21} = (-1)^{2+1} \cdot 8 = -8 \,\|$ $a_{22} = 15 \to A_{22} = (-1)^{2+2} \cdot 1 = 1$ Then $\begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix} \to \begin{pmatrix} 15 & -2 \\ -8 & 1 \end{pmatrix}$.

[step2] The determinant of the initial matrix is $\det A = \begin{vmatrix} 1 & 8 \\ 2 & 15 \end{vmatrix} = 1 \cdot 15 - 8 \cdot 2 = -1$, then

$\begin{pmatrix} 15 & -2 \\ -8 & 1 \end{pmatrix} \to \begin{pmatrix} 15/-1 & -2/-1 \\ -8/-1 & 1/-1 \end{pmatrix} = \begin{pmatrix} -15 & 2 \\ 8 & -1 \end{pmatrix}$.

And finally, **[step3]** $\begin{pmatrix} -15 & 2 \\ 8 & -1 \end{pmatrix} \to \begin{pmatrix} -15 & 2 \\ 8 & -1 \end{pmatrix}^T = \begin{pmatrix} -15 & 8 \\ 2 & -1 \end{pmatrix} \equiv A^{-1}$. It's quite important to check our solution: $\begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix} \cdot \begin{pmatrix} -15 & 8 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -15 & 8 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix}$. Notice, according to the property **[B]**, it's enough to check only that $\begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix} \cdot \begin{pmatrix} -15 & 8 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or to check only that $\begin{pmatrix} -15 & 8 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 8 \\ 2 & 15 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. So there is no need to check both equalities (and similarly in any other case), which is a very convenient thing, it is especially convenient for big-size matrixes, because multiplication of such matrixes is quite a tedious task.

Example. Find the inverse matrix for $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 1 & 3 & 5 \end{pmatrix}$.

[step1]

$a_{11} = 1 \to A_{11} = (-1)^{1+1} \cdot \begin{vmatrix} 1 & 2 \\ 3 & 5 \end{vmatrix} = -1 \,\|\, a_{12} = 2 \to A_{12} = (-1)^{1+2} \cdot \begin{vmatrix} -1 & 2 \\ 1 & 5 \end{vmatrix} = 7 \,\|\, a_{13} = 3 \to A_{13} = (-1)^{1+3} \cdot \begin{vmatrix} -1 & 1 \\ 1 & 3 \end{vmatrix} = -4 \,\|$

$a_{21} = -1 \to A_{21} = (-1)^{2+1} \cdot \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} = -1 \,\|\, a_{22} = 1 \to A_{22} = (-1)^{2+2} \cdot \begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix} = 2 \,\|\, a_{23} = 2 \to A_{23} = (-1)^{2+3} \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = -1 \,\|$

$a_{31} = 1 \to A_{31} = (-1)^{3+1} \cdot \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 1 \,\|\, a_{32} = 3 \to A_{22} = (-1)^{3+2} \cdot \begin{vmatrix} 1 & 3 \\ -1 & 2 \end{vmatrix} = -5 \,\|\, a_{33} = 5 \to A_{33} = (-1)^{3+3} \cdot \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3 \,\|$

Then: $A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 1 & 2 \\ 1 & 3 & 5 \end{pmatrix} \to \begin{pmatrix} -1 & 7 & -4 \\ -1 & 2 & -1 \\ 1 & -5 & 3 \end{pmatrix}$.

[step2] The determinant of the initial matrix is $\det A=\begin{vmatrix}1 & 2 & 3\\ -1 & 1 & 2\\ 1 & 3 & 5\end{vmatrix}$. Let's calculate it (expand along the first row):

$$1\cdot(-1)^{1+1}\cdot\begin{vmatrix}1 & 2\\ 3 & 5\end{vmatrix}+2\cdot(-1)^{1+2}\cdot\begin{vmatrix}-1 & 2\\ 1 & 5\end{vmatrix}+3\cdot(-1)^{1+3}\cdot\begin{vmatrix}-1 & 1\\ 1 & 3\end{vmatrix}=\begin{vmatrix}1 & 2\\ 3 & 5\end{vmatrix}-2\cdot\begin{vmatrix}-1 & 2\\ 1 & 5\end{vmatrix}+3\cdot\begin{vmatrix}-1 & 1\\ 1 & 3\end{vmatrix}=-1-2\cdot(-7)+3\cdot(-4)=1,$$

then $\det A=1$ and $\begin{pmatrix}-1 & 7 & -4\\ -1 & 2 & -1\\ 1 & -5 & 3\end{pmatrix}\to\begin{pmatrix}-1/1 & 7/1 & -4/1\\ -1/1 & 2/1 & -1/1\\ 1/1 & -5/1 & 3/1\end{pmatrix}=\begin{pmatrix}-1 & 7 & -4\\ -1 & 2 & -1\\ 1 & -5 & 3\end{pmatrix}$.

And finally **[step3]** $\begin{pmatrix}-1 & 7 & -4\\ -1 & 2 & -1\\ 1 & -5 & 3\end{pmatrix}^{T}\to\begin{pmatrix}-1 & -1 & 1\\ 7 & 2 & -5\\ -4 & -1 & 3\end{pmatrix}=A^{-1}$, now we can check that $A^{-1}\cdot A=E$.

For Reader's practice:

[1] Solve the equation $\begin{vmatrix}\sin x & \cos x & \cos x\\ \cos x & \sin x & \cos x\\ \cos x & \cos x & \sin x\end{vmatrix}=0$ on the segment $-45^{o}\le x\le 45^{o}$. **Answer**: $x=45^{o}$.

[2] $P=\begin{pmatrix}\sqrt{3}/2 & 1/2\\ 1/2 & \sqrt{3}/2\end{pmatrix}$ and $A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ and $Q=PAP^{T}$. Find $P^{T}Q^{2005}P$. **Answer**: $\begin{pmatrix}1 & 2005\\ 0 & 1\end{pmatrix}$.

[3] Show that for any θ: $\begin{vmatrix}\sin\theta & \cos\theta & \sin 2\theta\\ \sin(\theta+120^{o}) & \cos(\theta+120^{o}) & \sin(2\theta+240^{o})\\ \sin(\theta-120^{o}) & \cos(\theta-120^{o}) & \sin(2\theta-240^{o})\end{vmatrix}=0$.

[4] Show that for any α,β,γ: $\begin{vmatrix}\sin\alpha & \cos\alpha & \sin(\alpha+\varphi)\\ \sin\beta & \cos\beta & \sin(\beta+\varphi)\\ \sin\gamma & \cos\gamma & \sin(\gamma+\varphi)\end{vmatrix}=0$.

[5] Calculate $\begin{vmatrix}1^2 & 2^2 & 3^2 & 4^2\\ 2^2 & 3^2 & 4^2 & 5^2\\ 3^2 & 4^2 & 5^2 & 6^2\\ 4^2 & 5^2 & 6^2 & 7^2\end{vmatrix}$ **Answer**: 0.

[6] Find the absolute value of the determinant $\begin{vmatrix}-1 & 2 & 1\\ 3+2\sqrt{2} & 2+2\sqrt{2} & 1\\ 3-2\sqrt{2} & 2-2\sqrt{2} & 1\end{vmatrix}$ **Answer**: $16\sqrt{2}$.

[7] Calculate $A=\begin{pmatrix}a & b & c\\ b & c & a\\ c & a & b\end{pmatrix}$ and $AA^{T}=1,\ abc=1$. Find the value of $a^3+b^3+c^3$ **Answer**: 4.

12

Complex numbers

Construction of complex numbers

Def. C is the field: C consists of real numbers R and the element i such that $i^2=-1$.
The addition and multiplication on C are extensions of the addition and multiplication on R, i.e., for any $a,b\in R\subset C$ we have:

$$a+b\ (\textit{addition by the rules of } C)=a+b\ (\textit{addition by the rules of } R),$$

$$a\cdot b\ (\textit{multiplication by the rules of } C)=a\cdot b\ (\textit{multiplication by the rules of } R).$$

C is called a field of complex numbers and elements of C are called complex numbers.

Let's build C. We fix the any set R of real numbers. Let's consider the set C_{aux} of all pairs $\{(a,b)\}$ where a,b are any real numbers. We define addition and multiplication on C_{aux}.

Def. $(a,b)\oplus(c,d)\equiv/\,by\ def\,/\equiv(a+c,b+d)$ and $(a,b)\bullet(c,d)\equiv(ac-bd,ad+bc)$.

How to remember it? Imagine that any pair (a,b) denotes the element $a+b\cdot i$ of some field, where $i^2=-1$. Then $(a,b)\oplus(c,d)\leftrightarrow(a+b\cdot i)+(c+d\cdot i)=(a+c)+(b+d)\cdot i\leftrightarrow(a+c,b+d)$ and $(a,b)\bullet(c,d)\leftrightarrow(a+b\cdot i)\cdot(c+d\cdot i)=(a\cdot c-b\cdot d)+(a\cdot d+b\cdot c)\cdot i\leftrightarrow(ac-bd,ad+bc)$.

Assertion1. The set $(C_{aux},\oplus,\ \bullet)$ is a field.

Proof. These properties can be checked straightly. We will give an another proof.

Let's consider the set X of 2×2 matrixes of real numbers $\mathrm{X}\equiv\left\{A=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\,\|\,a,b\in R\right\}$.

We can describe the set X as a set of all 2×2 matrixes which elements on the main diagonal are equal and elements on the secondary diagonal are opposite.

Auxiliary1. The set X is a field under the matrix addition and multiplication.
Really $\forall A,B\in\mathrm{X}$ we have

$$A=\begin{pmatrix}a & -b\\ b & a\end{pmatrix},\ B=\begin{pmatrix}c & -d\\ d & c\end{pmatrix}\Rightarrow A+B=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}+\begin{pmatrix}c & -d\\ d & c\end{pmatrix}=\begin{pmatrix}a+c & -(b+d)\\ b+d & a+c\end{pmatrix}$$

and $A\cdot B=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\cdot\begin{pmatrix}c & -d\\ d & c\end{pmatrix}=\begin{pmatrix}ac-bd & -ad-bc\\ bc+ad & -bd+ac\end{pmatrix}=\begin{pmatrix}ac-bd & -(bc+ad)\\ bc+ad & ac-bd\end{pmatrix}$.

So X is closed under the matrix addition and multiplication. As addition of matrixes is associative and commutative, it is also associative and commutative on X. The zero matrix $\mathrm{O}\equiv\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$ obviously belongs to X. For any matrix $A=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\in\mathrm{X}$, an opposite matrix $-A=\begin{pmatrix}-a & b\\ -b & -a\end{pmatrix}$ also belongs to X. Then X is a commutative group under addition.

Multiplication of any 2×2 matrixes is associative and also distributive over addition, in particular it is true on X. Then X is a ring.

Let's show that all non-zero elements of X form a commutative group with respect to multiplication. At first we need to show that $\mathrm{X}\setminus\mathrm{O}$ is closed under multiplication. Let's notice that for any

$A\in\mathrm{X}\setminus\mathrm{O}\Rightarrow A=\begin{pmatrix} a & -b \\ b & a\end{pmatrix}\neq\begin{pmatrix} 0 & 0 \\ 0 & 0\end{pmatrix}$ we have $\det A=a^2+b^2\neq 0$. And similarly, for any other matrix $B\in\mathrm{X}\setminus\mathrm{O}$ we have $\det B\neq 0$, then for $A\cdot B$ we have $\det A\cdot B=\det A\cdot\det B\neq 0$, in any case $A\cdot B$ looks like $\begin{pmatrix} p & -k \\ k & p\end{pmatrix}$ and $p^2+k^2\neq 0$, then one of the numbers p,k is not zero, therefore $A\cdot B$ is not a zero matrix $A\cdot B\neq\mathrm{O}\Rightarrow A\cdot B\in\mathrm{X}\setminus\mathrm{O}$. Next, multiplication is associative on $\mathrm{X}\setminus\mathrm{O}$, because multiplication of any matrixes is associative. Multiplication is also commutative on X, really, $\begin{pmatrix} a & -b \\ b & a\end{pmatrix}\cdot\begin{pmatrix} c & -d \\ d & c\end{pmatrix}=\begin{pmatrix} ac-bd & -(bc+ad) \\ bc+ad & ac-bd\end{pmatrix}$ and

$\begin{pmatrix} c & -d \\ d & c\end{pmatrix}\cdot\begin{pmatrix} a & -b \\ b & a\end{pmatrix}=\begin{pmatrix} ac-bd & -(bc+ad) \\ bc+ad & ac-bd\end{pmatrix}$.

Any matrix $A\in\mathrm{X}\setminus\mathrm{O}$ is a nonsingular matrix $\det A\neq 0$, then it has an inverse one A^{-1}, we have to check that $A^{-1}\in\mathrm{X}\setminus\mathrm{O}$. Let's take any

$A\in\mathrm{X}\setminus\mathrm{O}\Rightarrow A=\begin{pmatrix} a & -b \\ b & a\end{pmatrix}\Rightarrow A^{-1}=\dfrac{1}{a^2+b^2}\cdot\begin{pmatrix} a & b \\ -b & a\end{pmatrix}\in\mathrm{X}\setminus\mathrm{O}$, then $\mathrm{X}\setminus\mathrm{O}$ is a commutative group with respect to multiplication. And X is a field.

Next, the mapping $f:\mathrm{X}\to C_{aux}$ such that $f\begin{pmatrix} a & -b \\ b & a\end{pmatrix}\equiv(a,b)$ is one-to-one mapping.

For any matrix from X the mapping f compares a pair of real numbers which stay in it's first column. For any matrixes $A,B\in\mathrm{X}$ we have $f(A+B)=f(A)\oplus f(B)$ and $f(A\cdot B)=f(A)\bullet f(B)$ (here $\oplus$ and $\bullet$ are addition and multiplication of pairs on C_{aux}). These two properties of f are obvious, just look above at the places above where we calculated a product and a sum of matrixes from X. Therefore $(C,\oplus,\bullet)$ is a field.

Comment. The Reader can refer to the 1-st book “Construction of numbers, length and area”, there was the **theorem4** in the chapter “Groups, Rings, Fields”. We use this theorem now to conclude that C_{aux} is a field. We will use it again very soon (the next page) to conclude that C is a field.

Let's replace every pair $(a,0)\in C_{aux}$ by the number a. We will get the set C, which consists of real numbers R and some pairs of real numbers: $C=R\cup(\mathit{pairs}\ (a,b)\ \mathit{of\ real\ numbers\ where}\ b\neq 0)$.

The set C is called a set of complex numbers. Let's turn it into a field.
We need to define addition and multiplication on C.

We define one-to-one mapping $f: C_{aux} \to C$ such that: $(a,0) \in C_{aux} \Rightarrow f(a,0) \equiv /by\ def/ \equiv a$ and for any other pair $(a,b) \in C_{aux} \Rightarrow f(a,b) \equiv /by\ def/ \equiv (a,b)$.

Any elements $z,m \in C$ can be uniquely represented as $z = f(a,b),\ m = f(c,d)$, and we define:

$z + m = f(a,b) + f(c,d) \equiv /by\ def/ \equiv f((a,b) \oplus (c,d))$ and

$z \cdot m = f(a,b) \cdot f(c,d) \equiv /by\ def/ \equiv f((a,b) \bullet (c,d))$.

Then $f: C_{aux} \to C$ is the mapping, and for any $(a,b),(c,d) \in C_{aux}$ we have

$f((a,b) \oplus (c,d)) = f(a,b) + f(c,d)$ and $f((a,b) \bullet (c,d)) = f(a,b) \cdot f(c,d)$.

From here immediately follows that C is a field, and we also have $R \subset C$. (Theorem4, **Book1**, page40)

Assertion2. The addition and multiplication on C are extensions of the addition and multiplications on $R \subset C$.

Proof. Let's fix arbitrary $a,b \in R \subset C$, in order to find their sum/product in C we need to find their preimages in C_{aux}. Let's denote for a while $\tilde{+}, \tilde{\cdot}$ -the addition and multiplication by the rules of C. So, $a = f(a,0)$ and $b = f(b,0)$, then

$a \tilde{+} b\ (in\ C) = f(a,0) \tilde{+} f(b,0) = //by\ def// = f((a,0) \oplus (b,0)) = f(a+b,0) = a+b \in R$ and also

$a \tilde{\cdot} b\ (in\ C) = f(a,0) \tilde{\cdot} f(b,0) = //by\ def// = f((a,0) \bullet (b,0)) = f(a \cdot b,0) = a \cdot b \in R$.

Everything is proved. As the addition and multiplication on C are extensions of the addition and multiplication on $R \subset C$, it is appropriate to use the same symbols $+,\cdot$ to denote these operations.

Def. C contains the pair $(0,1)$, this pair is denoted by the symbol $i \equiv (0,1)$.

Assertion3. The element i has the property $i^2 = -1$. Every element $z \in C$ can be uniquely represented as $z = a + b \cdot i \,\|\, a,b \in R$.

Proof. Let's take $i \equiv (0,1) \in C$, then

$i \cdot i = f(0,1) \cdot f(0,1) = //by\ def// = f((0,1) \bullet (0,1)) = f(-1,0) = -1$.

Existence of representation. Let $z \in C$. Then there exist the unique pair $(a,b) \in C_{aux}$ such that $f(a,b) = z$, for $(a,b) \in C_{aux}$ we have $(a,b) = (a,0) \oplus (0,b) = (a,0) \oplus ((b,0) \bullet (0,1))$ (in C_{aux}).

Then $z = f(a,b) = f((a,0) \oplus ((b,0) \bullet (0,1))) = f(a,0) + f((b,0) \bullet (0,1)) =$

$= f(a,0) + f(b,0) \cdot f(0,1) = a + b \cdot (0,1) = a + b \cdot i$.

Uniqueness of representation. Let $z = a + b \cdot i$ and $z = \bar{a} + \bar{b} \cdot i$, then

$a + b \cdot i = \bar{a} + \bar{b} \cdot i \Rightarrow a - \bar{a} = (\bar{b} - b) \cdot i$, if $\bar{b} = b$, then $a = \bar{a}$ and the representation is unique.

Let $\bar{b} \neq b$, then $a - \bar{a} = (\bar{b} - b) \cdot i \Rightarrow i = \dfrac{a - \bar{a}}{\bar{b} - b}$ there is a quotient of two real numbers on the right side, therefore the number on the right side belongs to R, then $i \in R$. But we have the equality $i^2 = -1$, which is impossible in R. Really, R is an ordered field, and for any element $r \in R$ there

must be $r^2 \geq 0$, so we have a contradiction, this contradiction came from the assumption $\bar{b} \neq b$. Then $\bar{b} = b$ and the representation is unique.

The assertion3 has a great practical importance. As any complex number z can be uniquely written as $a+b\cdot i$, this form of complex numbers is used in practice. Because it allows us to perform addition/subtraction/multiplication/division of complex numbers much faster. Really, the field of complex numbers C includes real numbers and pairs of real numbers, and in order to add or multiply these elements we should initially refer to the field C_{aux}, so it's not a fast and a comfortable way to operate in C. The notation $a+b\cdot i$ allows us to perform all arithmetical operations right in the field C, without referring to C_{aux}. Now we can discard C_{aux}, it was an intermediate auxiliary field, and we don't need it anymore.

Example1. [A] $(5+2i)+(6+7i)=11+9i$ and
$(5+2i)\cdot(6+7i)=5\cdot 6+5\cdot 7i+2i\cdot 6+2i\cdot 7i=30+47i+14i^2=30+47i-14=16+47i$.

[B] $(2+i)^3=(2+i)\cdot(2+i)\cdot(2+i)=(4+4i+i^2)\cdot(2+i)=(3+4i)\cdot(2+i)=6+3i+8i+4i^2=2+11i$.

[C] $\dfrac{1+i}{1-i}=\dfrac{(1+i)\cdot(1+i)}{(1-i)\cdot(1+i)}=\dfrac{1+i+i+i^2}{1-i^2}=\dfrac{2i}{2}=i$.

Two complex numbers $a+bi$ and $c+di$ are equal $\Leftrightarrow$ $a=c$ and $b=d$.
(it immediately follows from the uniqueness of representation).

Def. For any complex number $z=a+bi$ the real number a is called a real part of z and b is called an imaginary part of z. And we can write $a=\operatorname{Re} z$ and $b=\operatorname{Im} z$.

Assertion4. The field C is not an ordered field, i.e., there is no way to introduce any order relation ">" on C.

Auxiliary2. The ring of integer numbers Z can be ordered in the only one way, i.e., the order on Z, where all natural numbers are positive is the only possible order on Z.

Proof. Let's assume that Z is ordered in some other way, according to the simpliest properties of ordered rings/fields: $\forall a \in Z \,||\, a \neq 0 \Rightarrow a^2 > 0$, let's take $1 \in Z \Rightarrow 1^2 = 1 > 0$ so the number 1 is positive, any natural number can be represented as a sum of ones, then $\forall n \in N \in Z$ we have $n = 1+1+....+1 > 0$, then any natural number in Z must be positive. In any ordered ring, the element $-a$, which is opposite to some positive element a, must be negative, then every number from $-N \in Z$ is negative. The only number that we haven't considered yet is $0 \in Z$, but a zero element is not positive or negative in any ordered ring/field. We have deduced that only natural numbers N are positive in Z, therefore our order relation on Z is exactly the same as a standard order relation on Z.

Let's assume now that C is an ordered field, there exist some order relation ">" on C.
From the exercise6 (chapter "Groups, rings, fields" – Book I) follows that any subring/subfield of C is also an ordered ring/field. In particular, $Z \in C$ is an ordered ring, but the order relation on Z

is unique. We have $1>0$ $(in\ Z) \Leftrightarrow 1>0$ $(in\ C)$ (**exercise6**), then $-1<0$ $(in\ C)$.

If C is an ordered field, then for any $z \in C$ there must be $z^2 \geq 0$, in particular for $i \in C$ there must be $i^2 \geq 0$, but $i^2 = -1 < 0$, and we have a contradiction. So C is not an ordered field.

We know that in any **ordered** ring/field an absolute value (or module) $|a|$ of any element a is defined, this definition is made based on some order relation ">". By using the notion of absolute value we defined a limit of a sequence and exlored it's properties . All this theory can't be applied to C, because C is not an ordered field, so the old definition of $|a|$ can't be introduced in C. Anyway, it's very important to introduce the similar notion in C.

Def. For any complex number $z = a + bi$, the module (or absolute value) of z is $|z| \equiv \sqrt{a^2 + b^2}$.

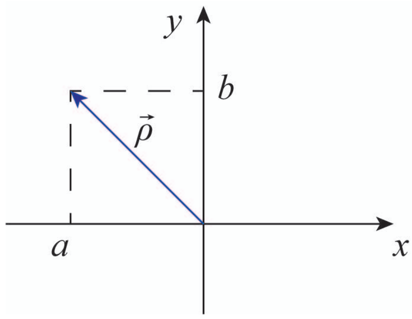
Exercise1. Show that **[A]** $\forall z \in C \Rightarrow |z| \geq 0$ and $|z| = 0 \Leftrightarrow z = 0$,
[B] $\forall z_1, z_2 \in C \Rightarrow |z_1 + z_2| \leq |z_1| + |z_2|$ and $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.

Def. For any complex number $z = a + bi$, the complex number $\bar{z} = a - bi$ is called a conjugate number.

Exercise2. Show that **[A]** $\forall z \in C \Rightarrow z \cdot \bar{z} = |z|^2$ **[B]** $|z| = |\bar{z}|$
[C] $\forall z_1, z_2 \in C \Rightarrow \overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2 \ \ || \ \ \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2 \ \ || \ \ \overline{z_1 / z_2} = \bar{z}_1 \cdot \bar{z}_2$.

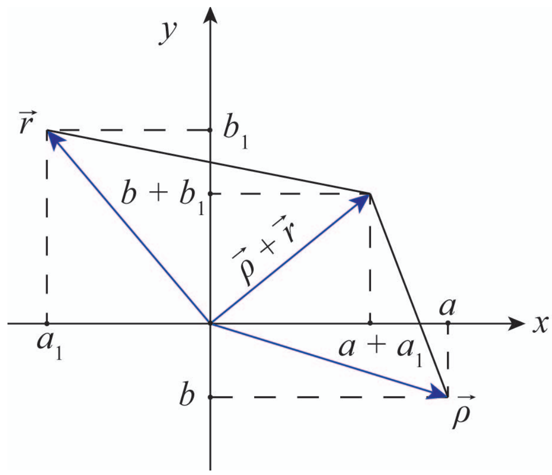
Geometrical representation of complex numbers

Any complex number $z = a + bi$ can be depicted as a radius vector $\vec{\rho}$ with coordinates (a,b) **[pict1]**. The axis Ox is called a real axis and Oy is called an imaginary axis. Such representation of complex numbers (as radius vectors) has great advantages. Addition of complex numbers can be represented as addition of radius vectors which depict these numbers **[pict2]**.
A module $|z|$ of any complex number z can be understanded as a length ρ of a radius vector $\vec{\rho}$ which depicts z.

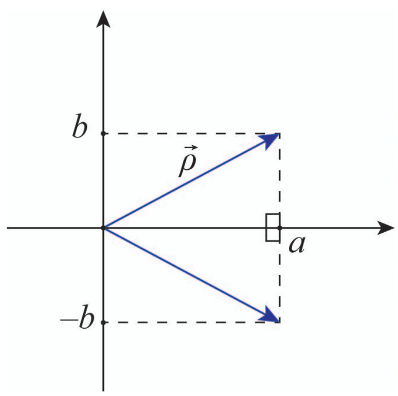


pict.1

For any complex number, which is depicted as a radius vector, the conjugate complex number can be depicted as a symmetrical radius vector **[pict3]** (with respect to Ox).
Other advantages of geometrical representation will be described later, when we introduce the next theory.



pict.2

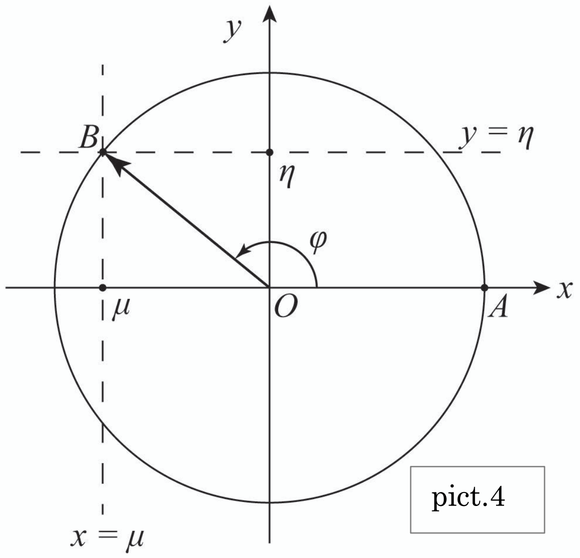


pict.3

Assertion5. For any non-zero complex number $z = a + bi$ there exist the unique angle $\varphi \in [0^o, 360^o)$ such that $z = |z| \cdot (\cos\varphi + i\sin\varphi)$. The angle φ is called an argument of z, and we can denote $\varphi = \arg z$.

Auxiliary3. For any pair of real numbers μ, η such that $\mu^2 + \eta^2 = 1$ there exist the angle $\varphi \in [0^o, 360^o)$ such that $\cos\varphi = \mu$, $\sin\varphi = \eta$.

Proof. **Existence**. Let's take some coordinate system Oxy. We draw the unit circle and the lines $x = \mu$ and $y = \eta$ **[pict4]**, these lines intersect at the point B with coordinates (μ, η), as $\mu^2 + \eta^2 = 1$, then B lies on the unit circle. The angle $\varphi \equiv \angle AOB$, which is counted from OA in the counterclockwise direction, is obviously the angle we need.



pict.4

Let's fix now any non-zero complex number $z \in C$,

then $z = a + bi = \sqrt{a^2 + b^2} \cdot \left(\frac{a}{\sqrt{a^2 + b^2}} + \frac{b}{\sqrt{a^2 + b^2}} \cdot i \right)$.

For numbers $\frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}}$ we obviously have

$\left(\frac{a}{\sqrt{a^2 + b^2}} \right)^2 + \left(\frac{b}{\sqrt{a^2 + b^2}} \right)^2 = 1$, then (auxiliary3) there exist $\varphi \in [0^o, 360^o)$ such that

$\cos\varphi = \frac{a}{\sqrt{a^2 + b^2}}$, $\sin\varphi = \frac{b}{\sqrt{a^2 + b^2}}$, then $z = \sqrt{a^2 + b^2} \cdot (\cos\varphi + i\sin\varphi) = |z| \cdot (\cos\varphi + i\sin\varphi)$ - the representation we need.

Let's show that such angle φ is unique. Let's assume that there exist some other angle $\alpha \in [0^o, 360^o)$, $\alpha \neq \varphi$, for which we also have $z = |z| \cdot (\cos\alpha + i\sin\alpha)$. Then $|z| \cdot (\cos\alpha + i\sin\alpha) = |z| \cdot (\cos\varphi + i\sin\varphi) \Rightarrow \cos\alpha + i\sin\alpha = \cos\varphi + i\sin\varphi$, as complex numbers are equal, their real and imaginary parts are equal, then $\cos\alpha = \cos\varphi$ and $\sin\alpha = \sin\varphi$.

As we have $\alpha \neq \varphi$, then $\alpha - \varphi \neq 0^o$, then $\cos(\alpha - \varphi) \neq 1$.

But $\cos(\alpha - \varphi) = \cos\alpha\cos\varphi + \sin\alpha\sin\varphi = \cos^2\varphi + \sin^2\varphi = 1$, we have a contradiction.

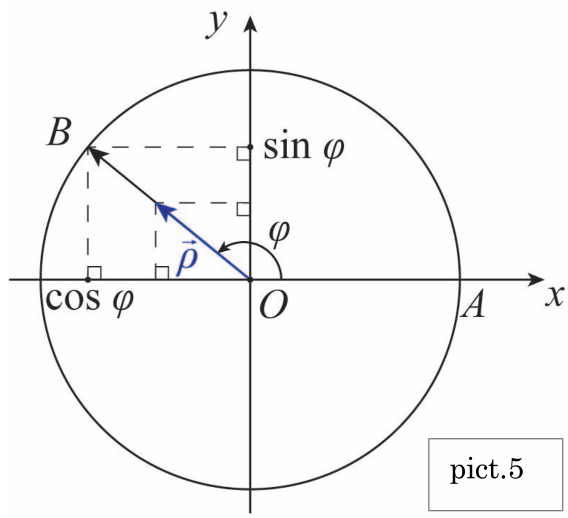
Then φ is unique.

Def. Quite often we use the letter ρ to denote $|z|$, so $\rho \equiv |z|$ and, as we proved above, for any non-zero complex number z there exist the unique representation $z = \rho \cdot (\cos\varphi + i\sin\varphi)$ **[T]**, where $\varphi \in [0^o, 360^o)$. **[T]** is called a trigonometric representation of z, or a trigonometric form of z.

Assertion6. When some non-zero complex number $z = a + bi$ is depicted as a radius vector $\vec{\rho}$ with coordinates (a,b), the angle φ from the trigonometric form **[T]** is the angle between $\vec{\rho}$ and Ox, which is counted from Ox in the counterclockwise direction.

Proof. From **[T]** follows that $z = a + bi = \rho \cdot (\cos\varphi + i\sin\varphi) \Rightarrow a = \rho\cos\varphi$ and $b = \rho\sin\varphi$.

The radius vector $\vec{\rho}$ has coordinates $(\rho\cos\varphi, \rho\sin\varphi)$ **[pict5]**. Let's draw the unit circle and build the angle $\angle AOB = \varphi$, which is counted in the counterclockwise direction from Ox. Then $\overrightarrow{OB}$ has coordinates $(\cos\varphi, \sin\varphi)$, then $\rho \cdot \overrightarrow{OB}$ must have the same coordinates $(\rho\cos\varphi, \rho\sin\varphi)$ as $\vec{\rho}$. Then $\rho \cdot \overrightarrow{OB}$ coincides with $\vec{\rho}$. (really, as coordinates of these vectors are equal, vectors are equal, and we have two equal vectors with the common start point O, so these vectors must coincide). And $\rho \cdot \overrightarrow{OB}$ for sure forms the angle φ with Ox, then the same is true for $\vec{\rho}$.



pict.5

Comment. The number $0 \in C$ can be also written in a trigonometric form. Let's find this form. Suppose we have a representation $0 = \rho \cdot (\cos\varphi + i\sin\varphi)$. Obviously, $0 = 0 + 0 \cdot i$, then $\rho \cdot (\cos\varphi + i\sin\varphi) = 0 + 0 \cdot i$. When complex numbers are equal, their real and imaginary parts are equal, then $\rho \cdot \cos\varphi = 0,\ \ \rho \cdot \sin\varphi = 0$. Then

$(\rho \cdot \cos\varphi)^2 = 0,\ \ (\rho \cdot \sin\varphi)^2 = 0 \Rightarrow \rho^2 \cdot \cos^2\varphi + \rho^2 \cdot \sin^2\varphi = 0 \Rightarrow \rho^2(\cos^2\varphi + \sin^2\varphi) = 0 \Rightarrow$

$\Rightarrow \rho^2 \cdot 1 = 0 \Rightarrow \rho = 0$. So, in the representation $0 = \rho \cdot (\cos\varphi + i\sin\varphi)$ there must be $\rho = 0$. From here immediately follows that any angle φ is appropriate. So, let's make an agreement that φ is any angle from $[0^o, 360^o)$.

Assertion7. $z_1 = \rho_1(\cos\varphi_1 + i\sin\varphi_1)$ and $z_2 = \rho_2(\cos\varphi_2 + i\sin\varphi_2)$ are two non-zero complex numbers. Then $z_1 \cdot z_2 = \rho_1\rho_2 \cdot \left(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\right)$ and

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} \cdot (\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)).$$

And the geometrical meaning is: when we multiply two complex numbers, we must multiply their modules and add their arguments. When we divide two complex numbers, we must divide their modules and subtract one argument from another.

Proof. $z_1 \cdot z_2 = \rho_1\rho_2 \cdot (\cos\varphi_1 + i\sin\varphi_1) \cdot (\cos\varphi_2 + i\sin\varphi_2) =$

$= \rho_1\rho_2 \cdot \left([\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2] + i[\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2]\right) =$

$= \rho_1\rho_2 \cdot \left(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\right)$. By the way, from here immediately follows that for any complex number $z = \rho(\cos\varphi + i\sin\varphi)$ and for any natural number $n \in \mathrm{N}$ we have

$z^n = \rho^n(\cos n\varphi + i\sin n\varphi)$ **[De Moivre's formula]**.

Next, $\dfrac{z_1}{z_2}=\dfrac{\rho_1(\cos\varphi_1+i\sin\varphi_1)}{\rho_2(\cos\varphi_2+i\sin\varphi_2)}=\dfrac{\rho_1(\cos\varphi_1+i\sin\varphi_1)}{\rho_2(\cos\varphi_2+i\sin\varphi_2)}\cdot\dfrac{(\cos\varphi_2-i\sin\varphi_2)}{(\cos\varphi_2-i\sin\varphi_2)}=$

$$=\frac{\rho_1}{\rho_2}\cdot\frac{[\cos\varphi_1\cos\varphi_2+\sin\varphi_1\sin\varphi_2]+i[\sin\varphi_1\cos\varphi_2-\cos\varphi_1\sin\varphi_2]}{(\cos\varphi_2)^2+(\sin\varphi_2)^2}=\frac{\rho_1}{\rho_2}\cdot(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)).$$

Example2. Calculate $(\cos 72^o+i\sin 72^o)^{40}$. **Solution:** Let's denote $z=\cos 72^o+i\sin 72^o=1\cdot(\cos 72^o+i\sin 72^o)$, then

$(\cos 72^o+i\sin 72^o)^{40}=z^{40}=1^{40}\cdot(\cos(40\cdot 72^o)+i\sin(40\cdot 72^o))=\cos(8\cdot 360^o)+i\sin(8\cdot 360^o)=1.$

Example3. Calculate $\left(1+\sqrt{3}i\right)^{2018}$. **Solution**: Let's write this number in a trigonometric form and use **[De Moivre's formula]** $(\rho(\cos\varphi+i\sin\varphi))^n=\rho^n(\cos n\varphi+i\sin n\varphi)$.

We have the number $\left(1+\sqrt{3}i\right)^{2018}$.

Let's represent $1+\sqrt{3}i\equiv\rho(\cos\varphi+i\sin\varphi)$, there must be $\rho=\sqrt{1^2+(\sqrt{3})^2}=\sqrt{1+3}=2$ and

$\cos\varphi=\dfrac{1}{2},\ \sin\varphi=\dfrac{\sqrt{3}}{2}\Rightarrow\cos\varphi=\dfrac{1}{2},\ \sin\varphi=\dfrac{\sqrt{3}}{2}\Rightarrow\varphi=60^o$, then $1+\sqrt{3}i\equiv 2(\cos 60^o+i\sin 60^o)$.

So $(1+\sqrt{3}i)^{2018}\equiv\left(2(\cos 60^o+i\sin 60^o)\right)^{2018}=2^{2018}\cdot\left(\cos(2018\cdot 60^o)+i\sin(2018\cdot 60^o)\right)=.$

$2^{2018}\cdot\left(\cos(336\cdot 360^o+120^o)+i\sin(336\cdot 360^o+120^o)\right)=2^{2018}\cdot\left(\cos(120^o)+i\cos(120^o)\right)=$

$$=2^{2018}\cdot\left(-\frac{1}{2}+\frac{\sqrt{3}}{2}i\right)=2^{2017}\cdot\left(-1+\sqrt{3}i\right).$$

Example4. Calculate $\left(\dfrac{1}{i-1}\right)^{100}$.

Solution: $\dfrac{1}{i-1}=\dfrac{1+i}{(1-i)(1+i)}=\dfrac{1+i}{2}=\dfrac{1}{2}(1+i)=\dfrac{1}{2}\cdot\sqrt{2}\cdot\left(\dfrac{1}{\sqrt{2}}+i\cdot\dfrac{1}{\sqrt{2}}\right)=\dfrac{1}{\sqrt{2}}(\cos 45^o+i\sin 45^o)$.

Then $\left(\dfrac{1}{i-1}\right)^{100}=\left(\dfrac{1}{\sqrt{2}}\right)^{100}\left(\cos(100\cdot 45^o)+i\sin(100\cdot 45^o)\right)=$

$$=\frac{1}{2^{50}}(\cos(12\cdot 360^o+180^o)+i\sin(12\cdot 360^o+180^o))=\frac{1}{2^{50}}(\cos(180^o)+i\sin(180^o))=-\frac{1}{2^{50}}.$$

Example5. $z_1,z_2,\dots z_n$ are complex numbers such that $|z_1|=|z_2|=\dots=|z_n|=1$.

And $z=\left(\sum_{k=1}^{n}z_k\right)\cdot\left(\sum_{k=1}^{n}\dfrac{1}{z_k}\right)$. Show that z is a real number and $0\le z\le n^2$.

Solution. For any z_k we have $|z_k|=1\Leftrightarrow z_k\cdot\bar{z}_k=1\Rightarrow\dfrac{1}{z_k}=\bar{z}_k$.

Therefore $z=\left(\sum_{k=1}^{n} z_k\right)\cdot\left(\sum_{k=1}^{n}\frac{1}{z_k}\right)=\left(\sum_{k=1}^{n} z_k\right)\cdot\left(\sum_{k=1}^{n}\bar{z}_k\right)=(z_1+z_2+\ldots+z_n)\cdot(\bar{z}_1+\bar{z}_2+\ldots+\bar{z}_n)=$

$=(z_1+z_2+\ldots+z_n)\cdot\overline{(z_1+z_2+\ldots+z_n)}=|z_1+z_2+\ldots+z_n|$ - it is a non-negative real number, then $z=|z_1+z_2+\ldots+z_n|\geq 0$. Also $z=|z_1+z_2+\ldots+z_n|\leq|z_1|+|z_2|+\ldots+|z_n|=n$. And we have $0\leq z\leq n$.

Example6. $z_1,z_2\ldots.\ z_n$ are complex numbers such that $|z_1|=|z_2|=\ldots=|z_n|=1$. Show that $\left|\sum_{k=1}^{n} z_k\right|=\left|\sum_{k=1}^{n}\frac{1}{z_k}\right|$.

Solution. For any z_k we have $|z_k|=1\Leftrightarrow z_k\cdot\bar{z}_k=1\Rightarrow z_k=\frac{1}{\bar{z}_k}$. Therefore $\left|\sum_{k=1}^{n} z_k\right|=\left|\sum_{k=1}^{n}\frac{1}{\bar{z}_k}\right|$,

as $\frac{\bar{z}}{\bar{h}}=\overline{\left(\frac{z}{h}\right)}\ \forall z,h\in C$, then $\frac{1}{\bar{z}_k}=\frac{\bar{1}}{\bar{z}_k}=\overline{\left(\frac{1}{z_k}\right)}$, then $\left|\sum_{k=1}^{n} z_k\right|=$

$$=\left|\sum_{k=1}^{n}\frac{1}{\bar{z}_k}\right|=\left|\sum_{k=1}^{n}\overline{\left(\frac{1}{z_k}\right)}\right|=\left|\overline{\left(\frac{1}{z_1}\right)}+\overline{\left(\frac{1}{z_2}\right)}+\ldots+\overline{\left(\frac{1}{z_n}\right)}\right|=\left|\overline{\left(\frac{1}{z_1}\right)+\left(\frac{1}{z_2}\right)+\ldots+\left(\frac{1}{z_n}\right)}\right|=\begin{bmatrix}|z|=|\bar{z}|\\ \forall z\in C\end{bmatrix}=$$

$=\left|\left(\frac{1}{z_1}\right)+\left(\frac{1}{z_2}\right)+\ldots+\left(\frac{1}{z_n}\right)\right|=\left|\sum_{k=1}^{n}\frac{1}{z_k}\right|$, everything is proved.

Assertion8. There are no zero divisors in C $\Leftrightarrow$ if $z_1\cdot z_2=0$ in C, then $z_1=0$ or $z_2=0$.

Comment. We have shown earlier that there are no zero divisors in any ordered ring/field. That's why we didn't need to check the similar assertion for Z,Q,R.

Proof. Let's assume that our assertion is not true, it means that there exist some **non-zero** numbers z_1,z_2 such that $z_1\cdot z_2=0$. Any non-zero complex number can be uniquely represented in a trigonometric form. So $z_1=\rho_1(\cos\varphi_1+i\sin\varphi_1)$ and $z_2=\rho_2(\cos\varphi_2+i\sin\varphi_2)$, then $z_1\cdot z_2=\rho_1\rho_2\cdot(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2))=0$. From here it's very easy to get the equality $\rho_1\rho_2=0$. Then $\rho_1=0$ or $\rho_2=0$, it means that $z_1=0$ or $z_2=0$, and we have a contradiction. So, there are no zero divisors in C.

Def. Let $z\in C$ is any complex number and $n\in\mathrm{N}$. If there exist some $\mu\in C$ such that $(\mu)^n=z$ then μ is called an $n-th$ root of z, and we denote it $\mu\equiv\sqrt[n]{z}$.

Assertion9. For $0\in C$ and for any $n\in\mathrm{N}$ we have $\sqrt[n]{0}=0$. And for any non-zero complex number $z\in C$ and any $n\in\mathrm{N}$ there exist exactly n different complex numbers $\mu_0,\mu_1\ldots\ \mu_{n-1}\in C$ such that $(\mu_0)^n=z,\ (\mu_1)^n=z\ \ldots.\ (\mu_{n-1})^n=z$ (the symbol $\sqrt[n]{z}$ has n different values $\mu_0,\mu_1\ldots\ \mu_{n-1}$).

Proof. Let $0 \in C$. Let's fix any $n \in \mathrm{N}$, let $(\mu)^n = 0 \Leftrightarrow \mu \cdot \mu \cdot \ldots \cdot \mu = 0$. From the assertion8 follows that one of the factors $\mu \cdot \mu \cdot \ldots \cdot \mu$ is zero, then $\mu = 0$.

Let $z \in C$ is any non-zero complex number. Then it can be uniquely represented as $z = \rho(\cos\varphi + i\sin\varphi)$. Let there exist some complex number $\mu \in C$ such that $\mu^n = z$. If $\mu = 0$, then $z = 0$, which is not true, then $\mu \neq 0$ and μ can be uniquely represented as $\mu = l(\cos\theta + i\sin\theta)$. From the equality $\mu^n = z$ and **[De Moivre's formula]** follows that $l^n(\cos n\theta + i\sin n\theta) = \rho(\cos\varphi + i\sin\varphi)$ -these complex numbers are equal, therefore their real and imaginary parts are equal: $l^n \cos n\theta = \rho\cos\varphi$ and $l^n \sin n\theta = \rho\sin\varphi$ **[M]** from here follows $\left(l^n \cos n\varphi\right)^2 = (\rho\cos\theta)^2$ and $\left(l^n \sin n\varphi\right)^2 = (\rho\sin\theta)^2$ $\Leftrightarrow l^{2n}\cos^2 n\varphi = \rho^2\cos^2\theta$ and $l^{2n}\sin^2 n\varphi = \rho^2\sin^2\theta$, let's add these equalities $l^{2n}(\cos^2 n\varphi + \sin^2 n\varphi) = \rho^2(\cos^2\theta + \sin^2\theta) \Leftrightarrow l^{2n} = \rho^2$, as ρ, l are positive real numbers, from $l^{2n} = \rho^2$ follows that $\rho = l^n$, then $l = \sqrt[n]{\rho}$. Let's return to **[M]**, so $(\sqrt[n]{\rho})^n \cos n\theta = \rho\cos\varphi$ and $(\sqrt[n]{\rho})^n \sin n\theta = \rho\sin\varphi$, then $\cos n\theta = \cos\varphi$ and $\sin n\theta = \sin\varphi$ **[V]**.

Here φ is a fixed angle, and θ is the angle we need to find. From **[V]** we see that cosines and sines of $n\theta$ and φ are equal, then these angles define the same point B on the unit circle. We have the restrictions: $\theta \in [0^o, 360^o)$ and $\varphi \in [0^o, 360^o)$. Then $n\theta$ must belong to one of the next sets $[0^o,\ 360^o),\ [360^o,\ 360^o \cdot 2),\ [360^o \cdot 2,\ 360^o \cdot 3)\ \ldots\ [360^o \cdot (n-1),\ 360^o \cdot n)$, and φ belongs to $[0^o, 360^o)$. As $n\theta$ and φ define the same point on the unit circle, only the next variants are possible: $n\theta = \varphi,\ n\theta = \varphi + 360^o,\ n\theta = \varphi + 360^o \cdot 2\ ,\ n\theta = \varphi + 360^o \cdot 3 \ldots\ldots n\theta = \varphi + 360^o \cdot (n-1)$ it is equivalent to $\theta = \dfrac{\varphi}{n},\ \theta = \dfrac{\varphi + 360^o}{n},\ \theta = \dfrac{\varphi + 360^o \cdot 2}{n}\ ,\ \theta = \dfrac{\varphi + 360^o \cdot 3}{n}\ \ldots\ldots \theta = \dfrac{\varphi + 360^o \cdot (n-1)}{n}$.

So, we have found $l = \sqrt[n]{\rho}$ and we have also found all the possible values of θ, they define exactly n different complex numbers:

$$\left\{\mu_k = \sqrt[n]{\rho} \cdot \left(\cos\left(\frac{\varphi + 360^o \cdot k}{n}\right) + i \cdot \sin\left(\frac{\varphi + 360^o \cdot k}{n}\right)\right) \,\|\, k = 0,1,2\ \ldots\ (n-1)\right\}.$$

Now we need to check that all these numbers μ_k really satisfy the equation $\mu^n = z$, it's very easy to do by using the assertion7. Everything is proved.

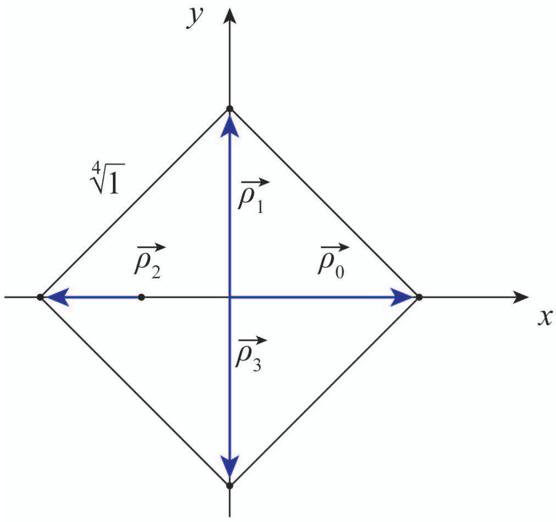
During the proof we have deduced a very important formula, the explicit value of n-th root.

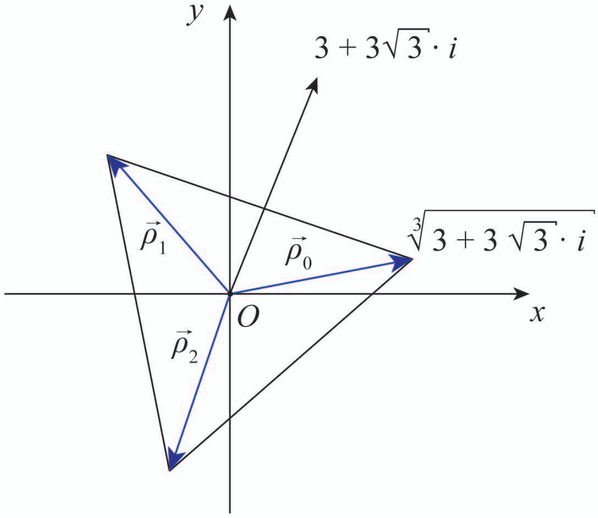
If $z=\rho(\cos\varphi+i\sin\varphi)$, then

$$\left\{\sqrt[n]{z}=\sqrt[n]{\rho}\cdot\left(\cos\left(\frac{\varphi+360^o\cdot k}{n}\right)+i\sin\left(\frac{\varphi+360^o\cdot k}{n}\right)\right) \| k=0,1,2\ \ldots\ (n-1)\right\}.$$

And the geometrical meaning is: for any non-zero complex number z all the values of $\sqrt[n]{z}$ can be depicted as radius vectors, which end points are vertexes of some regular polygon with n sides **[pict6]**, the center of such polygon is always located at the origin.

pict.6





Example7. Calculate $\sqrt[4]{256}$.

Solution. We will use the formula

$$\sqrt[n]{z}=\sqrt[n]{|z|}\left(\cos\frac{\varphi+360^o\cdot k}{n}+i\sin\frac{\varphi+360^o\cdot k}{n}\right)\| k=0,1,2\ldots\ n-1.$$

In our case $z=256$ and $n=4$, so $\sqrt[n]{\rho}=\sqrt[4]{|256|}=4$.

Here $\varphi\equiv\arg z=0$, then we can write:

$$\sqrt[n]{256}=4\cdot\left(\cos\frac{360^o\cdot k}{4}+i\sin\frac{360^o\cdot k}{4}\right)\| k=0,1,2,3$$

For $k=0\Rightarrow 4\cdot\left(\cos\frac{360^o\cdot 0}{4}+i\sin\frac{360^o\cdot 0}{4}\right)=4$,

for $k=1\Rightarrow 4\cdot\left(\cos\frac{360^o\cdot 1}{4}+i\sin\frac{360^o\cdot 1}{4}\right)=4i$,

for $k=2\Rightarrow 2\cdot\left(\cos\frac{360^o\cdot 2}{4}+i\sin\frac{360^o\cdot 2}{4}\right)=-4$,

for $k=3\Rightarrow 2\cdot\left(\cos\frac{360^o\cdot 3}{4}+i\sin\frac{360^o\cdot 3}{4}\right)=-4i$.

Then $\sqrt[4]{256}$ has four different values: $4,\ 4i,\ -4,\ -4i$.

For Reader's practice:

[1] Calculate $\sum_{k=1}^{6}\left(\sin\frac{2\pi k}{7}-i\cos\frac{2\pi k}{7}\right)$ **Answer**: i.

[2] Calculate $4+5\cdot\left(-\frac{1}{2}+\frac{i\sqrt{3}}{2}\right)^{334}+3\cdot\left(-\frac{1}{2}+\frac{i\sqrt{3}}{2}\right)^{365}$ **Answer**: $\sqrt{3}i$.

[3] Show that the number $\left(\frac{\sqrt{3}}{2}+\frac{i}{2}\right)^5+\left(\frac{\sqrt{3}}{2}-\frac{i}{2}\right)^5$ is purely real.

[4] Solve the equation $z^2+|z|=0$ (Hint, use the representation $z=x+iy \| x,y\in R$).

Answer: $z=0, z=i, z=-i$.

[5] [Advanced] How many solutions does the equation $z^3+\bar{z}=|z|$ have? **Answer**: four solutions.

[6] [Advanced] Solve the equation $z^6=z+\bar{z}$. **Answer**: $0,\ \sqrt[5]{2},\ \frac{1}{2}\pm i\frac{\sqrt{3}}{2},\ -\frac{\sqrt[5]{27}}{2}\pm\frac{\sqrt[10]{3}}{2}i$.

Def. For any complex number $z_0\in Z$ and any $\varepsilon>0$ the set of complex numbers $O_\varepsilon(z_0)\equiv\{z\,||\,|z-z_0|<\varepsilon\}$ is called an ε-neighborhood of z_0.

Any sequence $\{z_n\}$ of complex numbers uniquely defines two sequnces of real numbers $\{x_n\}$ and $\{y_n\}$, where $z_n=x_n+iy_n$ and conversely. The point z_0 is called a limit of the sequence $\{z_n\}\in Z$ if any neighborhood of z_0 contains all the terms of $\{z_n\}$, starting from some number.

Exercise3. Show that $\{z_n\}\to z_0$ (where $z_n=x_n+iy_n$ and $z_0=a+bi$)$\Leftrightarrow\{x_n\}\to a$ and $\{y_n\}\to b$. In other words, a sequence $\{z_n\}$ goes to z_0 if and only if the sequence of real parts of $\{z_n\}$ goes to the real part of z_0 and the sequence of imaginary parts of $\{z_n\}$ goes to the imaginary part of z_0.

A sequence of complex numbers $\{z_n\}$ is called fundamental, if for any (small) $\varepsilon>0$ there exist k such that $\forall m,n>k\Rightarrow|z_m-z_n|<\varepsilon$.

Exercise4. Show that: $\{z_n\}$ is fundamental (where $z_n=x_n+iy_n$) $\Leftrightarrow$ both $\{x_n\}$ and $\{y_n\}$ are fundamental.

Assertion10. The field of complex numbers C is a complete field: any fundamental sequence $\{z_n\}\in C$ converges to some limit $z_0\in C$.

Proof. Let's fix an arbitrary fundamental sequence $\{z_n\}$, we have $z_n=x_n+iy_n\,||\,\forall n$, then (exercise4) $\{x_n\}$ and $\{y_n\}$ are both fundamental, as R is a complete field, both these sequences converge: $\{x_n\}\to a,\ \{y_n\}\to b$, then (exercise3) $\{z_n\}$ converges to $a+bi$.

13

Analysis

Limits

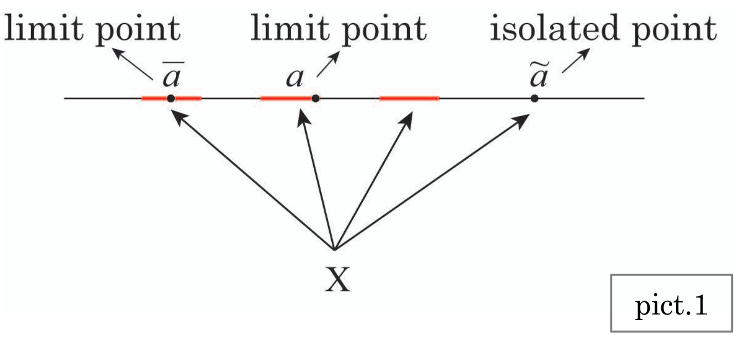
Earlier we have defined **3** types of points: internal points, boundary points, external points.

If some set $X \subset R$ is fixed, then **any** point $a \in R$ is an internal point of X, or a boundary point of X, or an external point of X.

It's very important to define the next **2** types of points: limit points and isolated points.

Def. Let $X \subset R$ is a set. And $a \in R$ is some real number (any real number can be called a point). If there exist some sequence $\{x_n\} \subset X \| x_n \neq a \ \forall n$ which goes to a, then a is called a **limit point** of X. If $\tilde{a} \in X$ and **there is no** any sequence $\{x_n\} \subset X \| x_n \neq \tilde{a} \ \forall n$ which goes to $\tilde{a}$, then $\tilde{a}$ is called an **isolated point** of X **[pict1].**

A limit point of X mustn't belong to X.
For example, for any half interval $(a,b]$, a is a limit point and it does not belong to $(a,b]$.
In the same time for any segment $[a,b]$, a is a limit point, and now it belongs to $[a,b]$.
Any isolated point of X belongs to X by definition.



pict.1

Exercise1. For any set $X \subset R$, any point $a \in X$ is a limit point of X **or** an isolated point of X, and no other variants.

Exercise2. Any internal point of X is a limit point of X. A boundary point of X may be a limit point of X and may not be a limit point of X. An external point of X is never a limit point of X.

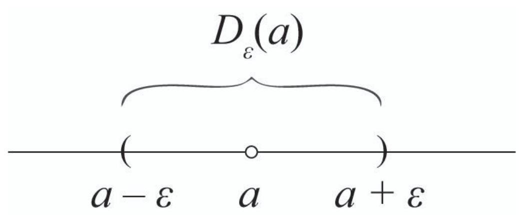
In order to define the first **3** types of points (internal, boundary, external) we used the notion of a neighborhood $O_\varepsilon(a)$. For the new **2** types of points it's much more convenient to use the notion of a deleted neighborhood $D_\varepsilon(a)$, in fact it is a neighborhood $O_\varepsilon(a)$ where the central point a is deleted.

Def. For any $a \in R$ and any $\varepsilon > 0$, the set of numbers $D_\varepsilon(a) \equiv O_\varepsilon(a) \setminus \{a\} \equiv \{x \| 0 < |x-a| < \varepsilon\}$ **[pict2]** is called an ε-deleted neighborhood of a, or just a deleted neighborhood of a.
The set $D_\varepsilon(a)$ can be depicted on the line as an interval without it's middle point.

Exercise3. Show that: **[pict2]**

[A] a is an isolated point of $X \Leftrightarrow$ there exist $D_\varepsilon(a)$ which does not contain any points of X.

[B] a is a limit point of $X \Leftrightarrow$ any deleted neighborhood $D_\varepsilon(a)$ contains at least one point of X.



pict.2

Def [A]. X is a set of real numbers, $f(x)$ is defined on X. And a is a limit point of X (a may belong to X and may not belong to X, therefore f may be defined at a and may be not defined at a). If **for any** sequence $\{x_n\} \subset X$ such that $\{x_n\} \to a \,\|\, x_n \neq a \;\forall n$, the sequence $\{f(x_n)\}$ always goes to some fixed number A, then A is called a limit of f at the point a, and we write $\lim_{x \to a \| x \in X} f(x) = A$.

Once again. If **for any** $\{x_n\} \subset X \,\|\, x_n \neq a \;(\forall n) \,\|\, \{x_n\} \to a$ we have $\{f(x_n)\} \to A$, then A is a limit of f at a, and we write $\lim_{x \to a \| x \in X} f(x) = A$.

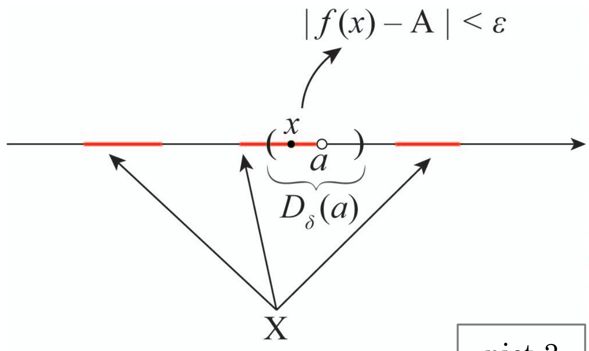
Notice: from this definition immediately follows that f may have only one limit at any point a, because the sequence $\{f(x_n)\}$ (as any other sequence) can't have two different limits $A \neq B$.

There is another equivalent definition.

Def [B]. X is a set of real numbers and $f(x)$ is defined on X. Let a is a limit point of X, and for any (small) positive $\varepsilon > 0$ there exist some positive $\delta > 0$, such that for any $x \in X : 0 < |x - a| < \delta$ we have $|f(x) - A| < \varepsilon$. Then A is called a limit of f at a, and we write $\lim_{x \to a \| x \in X} f(x) = A$.

Once again. If **for any** (small) positive $\varepsilon > 0$ there exist some positive $\delta > 0$ such that for any $x \in D_\delta(a) \cap X \Rightarrow |f(x) - A| < \varepsilon$, then $\lim_{x \to a \| x \in X} f(x) = A$.

Notice: the requirement "a is a limit point of X" is important, it guarantees that for any $\delta > 0$ the set $D_\delta(a) \cap X$ is not empty and it contains some points $x \in X$, for which the condition $|f(x) - A| < \varepsilon$ must be checked.



pict.3

We will show that this definition is equivalent to initial one, but let's explain at first it's meaning.

The meaning. Let $\lim_{x \to a \| x \in X} f(x) = A$, according to the **Def [B]**. **[pict3]**

Let's take a very small positive ε, for example $\varepsilon = 0{,}000001$, then there exist such neighborhood $D_\delta(a)$, that for any point $x \in D_\delta(a) \cap X$ the value $f(x)$ differs from A so slightly, that the "distance" between $f(x)$ and A is less than 0.000001.

So, in the set X everywhere near the point a, every value $f(x)$ is extremely close to A.

We can also say: when $x \in X$ approaches to a, the value $f(x)$ approaches "infinitely close" to A.

Really, let $x \in X$ approaches to a, then at some moment x will appear inside any fixed neighborhood $D_\delta(a)$, and when it happens, right away the value $f(x)$ is so close to A, that their difference is less than ε, where ε can be fixed from the very beginning as a very small positive number.

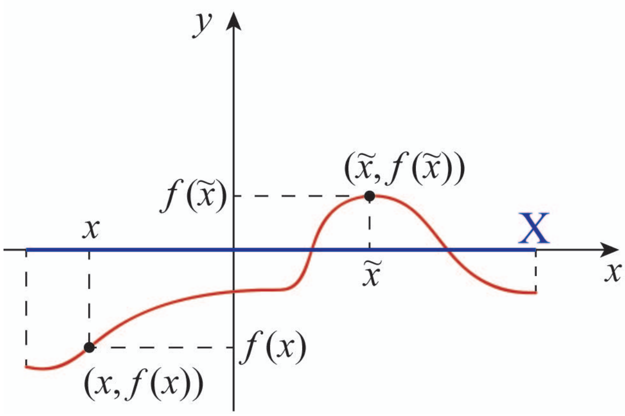
An important comment. In the limit definition (**Def [A]** or **Def [B]**) we do not allow x to reach the value a, we allow x to approach "infinitely close" to a, and we observe the value $f(x)$, does it approach infinitely close to some fixed number, or not. **But once and for all**: we forbid x to coincide with a, we allow x to approach infinitely close to a (and such approach is possible, because a is a limit point of X, and there exist $x \in \mathrm{X}$ so close to a as we want), but x can't coincide with a.

There are a lot of reasons for such restriction, without it we couldn't define a derivative of a function (it is one of the most important notions in mathematics). Let's explain briefly.

A derivative is defined as a limit of the function $\dfrac{f(x)-f(a)}{x-a}$ at the point a, and the function $\dfrac{f(x)-f(a)}{x-a}$ is not even defined at a. This function will be always defined on some deleted neighborhood of a. But if x reaches the value a, there appears an expression $\dfrac{0}{0}$, and such expression is not defined. So $\dfrac{f(x)-f(a)}{x-a}$ is not even defined at a, but it's limit at a is one of the central notions in mathematics. That's why the limit definition is made in such way that x never reaches the value a.

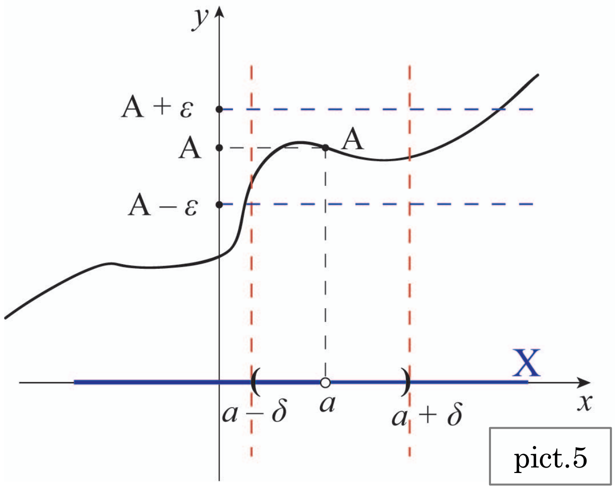
An important comment2. When we calculate $\lim_{x \to a || x \in \mathrm{X}} f(x) = \mathrm{A}$ the only values we need are the values of f on X. The function f may be defined at a, or may be not defined at a, it does not affect the existence of the limit $\lim_{x \to a || x \in \mathrm{X}} f(x)$ and it's value. And similarly, outside of the set X (on $R \setminus \mathrm{X}$) the function f may be not defined at all, or f may be defined at some points of $R \setminus \mathrm{X}$ and reach any values at these points, it does not affect the existence of $\lim_{x \to a || x \in \mathrm{X}} f(x)$ and it's value.

Def. X is a set of real numbers and f is defined on X. Let's fix any coordinate system Oxy. The set of points with coordinates $\{(x, f(x)) || x \in \mathrm{X}\}$ is called a graph of f on X, or just a graph of f **[pict4]**.



pict.4

Geometrical meaning of a limit. f has a limit A at a point a if for any small positive ε there exist some positive δ, such that the graph of f on the set $O_\delta(a)\cap X$ lies completely inside the strip which is bounded by the lines $y = A-\varepsilon,\ \ y = A+\varepsilon$ **[pict5]**.



pict.5

Let's show that definitions **Def [A]** and **Def [B]** are equivalent. Then in any concrete case we will be able to choose which definition to use.
Both definitions are really important, but the initial one **Def [A]** is very convenient for proving different properties of limits.

Proof. $\Rightarrow$Let **Def [A]** is true. **We want to show** that for any positive ε we can find some positive δ such that $\forall x \in D_\delta(a)\cap X \Rightarrow |f(x)-A|<\varepsilon$. Let's assume that it is not true, it means that there exist some concrete positive $\bar{\varepsilon}$, for which we can't find any appropriate δ. Therefore, for any $\delta>0$ there is always at least one "bad point" $\bar{x}$ in the set $D_\delta(a)\cap X$, for which we have $|f(\bar{x})-A|\ge\bar{\varepsilon}$.

Let's fix any decreasing sequence of positive numbers $\delta_1>\delta_2>\delta_3>\delta_4>....>0$ which goes to zero $\{\delta_n\}\to 0$, then we have the sequence of nested neighborhoods
$D_{\delta_1}(a)\supset D_{\delta_2}(a)\supset D_{\delta_3}(a)\supset D_{\delta_4}(a)$ ..... As we noticed above, in any neighborhood $D_{\delta_n}(a)$ there exist some "bad point" $\bar{x}_n\in D_{\delta_n}(a)\cap X$ such that $|f(\bar{x}_n)-A|\ge\bar{\varepsilon}$ **[T]**. Then we have the sequence $\{\bar{x}_n\}\subset X$ and $\{\bar{x}_n\}\to a \,||\, \bar{x}_n\ne a\ \forall n$. According to the **Def [A]**, for the sequence $\{f(\bar{x}_n)\}$ there must be $\{f(\bar{x}_n)\}\to A$, but we have $|f(\bar{x}_n)-A|\ge\bar{\varepsilon}\,||\,\forall n$, it means that $\bar{\varepsilon}$-neighborhood of A does not contain any terms of the sequence $\{f(\bar{x}_n)\}$, therefore A is not a limit of the sequence $\{f(\bar{x}_n)\}$. We have a contradiction. It proves that from **Def [A]** follows **Def [B]**.

Conversely $\Leftarrow$ Let **Def [B]** is true. **We want to show that** for any sequence
$\forall\{x_n\}\subset X\,||\,\{x_n\}\to a,\ \ x_n\ne a\ (\forall n)$: there must be $\{f(x_n)\}\to A$.
Let's fix any sequence $\forall\{x_n\}\subset X\,||\,\{x_n\}\to a,\ \ x_n\ne a\ (\forall n)$ and any positive $\varepsilon>0$,
according to the **Def [B]**, there exist $\delta>0$ such that $\forall x\in D_\delta(a)\cap X\Rightarrow|f(x)-A|<\varepsilon$.
As $\{x_n\}\subset X$ goes to a and $x_n\ne a\ (\forall n)$, the set $D_\delta(a)\cap X$ must contain all the terms of $\{x_n\}$, starting from some number k. And for any $x_n\in D_\delta(a)\cap X$ there must be
$|f(x_n)-A|<\varepsilon\Leftrightarrow f(x_n)\in O_\varepsilon(A)$. It means that ε neighborhood of A contains all the terms of $\{f(x_n)\}$, starting from the number k, then $\{f(x_n)\}\to A$. Everything is proved.

Sign conservation. Let $\lim_{x\to a||x\in X} f(x) = A$. If $A>0$, then there exist the deleted neighborhood $D_\delta(a)$ such that $\forall x \in D_\delta(a)\cap X$ the value $f(x)>0$. If $A<0$, then there exist the deleted neighborhood $D_\delta(a)$ such that $\forall x \in D_\delta(a)\cap X$ the value $f(x)<0$.

Put simply, if a limit of f at A is positive/negative, then all the values of f near A are also positive/negative.

Proof. Let $A>0$, let's fix for example $\varepsilon = A/2>0$, there exist $D_\delta(a)$ such that $\forall x \in D_\delta(a)\cap X$ we have $|f(x)-A|<A/2 \Leftrightarrow -A/2< f(x)-A<A/2 \Leftrightarrow A/2<f(x)<3A/2$, so the value $f(x)$ is "squeezed" between two positive numbers and therefore $f(x)$ is positive (for any $x\in D_\delta(a)\cap X$). Similarly, when A is a negative real number we can take $\varepsilon = |A|/2>0$.

Theorem1. Both limits $\lim_{x\to a||x\in X} f(x) = A$ and $\lim_{x\to a||x\in X} g(x) = B$ exist, then **[1]** $\lim_{x\to a||x\in X}(f(x)\pm g(x)) = A\pm B$ **[2]** $\lim_{x\to a||x\in X}(f(x)\cdot g(x)) = A\cdot B$ **[3]** $\lim_{x\to a||x\in X}\dfrac{f(x)}{g(x)} = \dfrac{A}{B}$ (if $g(x)\neq 0$ on X and $B\neq 0$) **[4]** $\lim_{x\to a||x\in X} \lambda f(x) = \lambda A$ for any real number $\lambda \in R$ **[5]** $\lim_{x\to a||x\in X}(f(x))^k = A^k$ for any $k\in N$.

An important comment. Notice that we require the existence of both limits $\lim_{x\to a|x\in X} f(x)$ and $\lim_{x\to a|x\in X} g(x)$, **only after that** we can guarantee that all the other limits **[1]-[5]** exist, but **NOT** conversely. We can't use any of the formulas **[1]-[5]** if we don't know for sure that both limits $\lim_{x\to a|x\in X} f(x)$ and $\lim_{x\to a|x\in X} g(x)$ do exist (there was a similar comment for sequence-limits).

Proof. The proof is very simple if we use the **Def [A]**. We have proved earlier quite similar theorems for sequences Basic properties of convergent sequences (Book1): if $\{x_n\}\to a$ and $\{y_n\}\to b$, then **[1.0]** $\{x_n+y_n\}\to a+b$, **[2.0]** $\{x_n\cdot y_n\}\to a\cdot b$, **[3.0]** $\left\{\dfrac{x_n}{y_n}\right\}\to \dfrac{a}{b}$ (if $y_n\neq 0\ \forall n$ and $b\neq 0$).

Let's fix an arbitrary sequence $\{x_n\}\subset X \,||\, x_n\neq a\ (\forall n) \,||\, \{x_n\}\to a$ from the condition $\lim_{x\to a|x\in X} f(x) = A$ follows that $\{f(x_n)\}\to A$ and from the condition $\lim_{x\to a|x\in X} f(x) = B$ follows that $\{g(x_n)\}\to B$. Then from **[1.0]**, **[2.0]**, **[3.0]** we have $\{f(x_n)\pm g(x_n)\}\to A\pm B$ and $\{f(x_n)\cdot g(x_n)\}\to A\cdot B$ and $\left\{\dfrac{f(x_n)}{g(x_n)}\right\}\to\dfrac{A}{B}$. From here follows that **[1] [2] [3]** are true.

Next, **[4]** immediately follows from **[2]** if we take a constant function $g(x)\equiv\lambda$ for any $x\in X$. And **[5]** can be easily derived from **[2]** (show how).

Squeeze theorem for functions. Functions f,g,h are defined on X and:
$\lim_{x\to a||x\in \mathrm{X}} f(x)=\mathrm{A}$ and $\lim_{x\to a||x\in \mathrm{X}} h(x)=\mathrm{A}$ and there exist some deleted neighborhood $D_\delta(a)$ such that $\forall x\in D_R(a)\cap \mathrm{X}:\ f(x)\le h(x)\le g(x)$, then also $\lim_{x\to a||x\in \mathrm{X}} h(x)=\mathrm{A}$.

Proof. Let's fix an arbitrary sequence $\{x_n\}\subset \mathrm{X}:\{x_n\}\to a$. From $\lim_{x\to a|x\in \mathrm{X}} f(x)=\mathrm{A}$ and $\lim_{x\to a|x\in \mathrm{X}} g(x)=\mathrm{A}$ follows that $\{f(x_n)\}\to \mathrm{A}$ and $\{g(x_n)\}\to \mathrm{A}$, then $f(x_n)\le h(x_n)\le g(x_n)$ and from the squeeze theorem for sequences we have $\{h(x_n)\}\to \mathrm{A}$. Everything is proved.

Exercise1 (Set replacement). $\lim_{x\to a||x\in \mathrm{X}} f(x)=\mathrm{A}$. Then for any set $\Omega\subset \mathrm{X}$, for which a is also a limit point, there must be $\lim_{x\to a||x\in \Omega} f(x)=\mathrm{A}$.

We have already provided the definition of a continuous function (it was done for the practical purposes in the book1, we needed it to understand for which $a\in R$ the value $\sqrt[k]{a}$ exists).

Def [A1]. f is defined on X and $a\in \mathrm{X}$. **If** a is a limit point of X and for any sequence $\{x_n\}\subset \mathrm{X}\,||\,x_n\ne a\ \forall n\,||\,\{x_n\}\to a$ we have $\{f(x_n)\}\to f(a)$, then we say "f is continuous at a".
If a is an isolated point of X, then f is continuous at a (by definition).

Notice, as any point $a\in \mathrm{X}$ is a limit point of X or an isolated point of X, then for any point $a\in \mathrm{X}$ it's possible to check, is f continuous at a or not.

When a is a limit point of X, continuity at a means that
[A] The limit $\lim_{x\to a||x\in \mathrm{X}} f(x)$ exists **[B]** This limit is equal to $f(a)$.

So, continuity is just a particular case of a limit-existence. But now f must be defined at a.
By using the limit definition **Def [B]** we can write an equivalent definition.

Def [B1]. f is defined on X and $a\in \mathrm{X}$. **If** a is a limit point of X and $\forall\varepsilon>0\ \exists\delta>0\,||\,\forall x\in \mathrm{X}:0<|x-a|<\delta\Rightarrow|f(x)-f(a)|<\varepsilon$, then "$f$ is continuous at a".
If a is an isolated point of X, then f is continuous at a (by definition).

This definition can be simplified, we can change the requirement $0<|x-a|<\delta$ to $|x-a|<\delta \Leftrightarrow x\in \mathrm{O}_\delta(a)$, really, by doing it we just include the case $x=a$, but then $|f(x)-f(a)|=|f(a)-f(a)|=0<\varepsilon$. So, when a is a limit point of X, we have the tantamount requirement $\forall\varepsilon>0\ \exists\delta>0\,||\,\forall x\in \mathrm{X}\cap \mathrm{O}_\delta(a)\Rightarrow|f(x)-f(a)|<\varepsilon$ **[R]**. Notice that the requirement **[R]** is obviously true for any isolated point $a\in \mathrm{X}$, really as a is an isolated point, there exist some $\mathrm{O}_{\bar\delta}(a)$ such that $\mathrm{O}_{\bar\delta}(a)$ has no common points with X except a.

Then for any positive $\varepsilon > 0$, we will always take $\bar{\delta} > 0$ and then the only $x \in O_{\bar{\delta}}(a) \cap X$ is $x = a$, for which the condition $|f(x) - f(a)| < \varepsilon$ is obviously true.

So, we can seriously abbreviate our definition **Def [B1]**.

The main definition of a continuous function [B]. f is defined on X and $a \in X$.

If $\forall \varepsilon > 0\ \exists \delta > 0 \,\|\, \forall x \in X \cap O_\delta(a) \Rightarrow |f(x) - f(a)| < \varepsilon$, then " f is continuous at a ".

(Show that this definition is (logically) equivalent to the **Def [B1]**, we have just shown that it follows from **Def [B1]**, and it's time to prove conversely). We can also abbreviate **Def [A1]** by throwing away the requirement $x_n \neq a$ $(\forall n)$ and the phrase about an isolated point a.

(Show that this definition is equivalent to **Def [A1]**).

The main definition of a continuous function [A]. f is defined on X and $a \in X$.

If $\forall \{x_n\} \subset X \,\|\, \{x_n\} \to a \Rightarrow \{f(x_n)\} \to f(a)$, then " f is continuous at a ".

By using **the main definition of a continuous function [B]** it's easy to prove the next theorem.

Theorem2. f, g are defined on X and $a \in X$. Both functions f, g are continuous at a, then:

[1] $f(x) \pm g(x)$ is continuous at a **[2]** $f(x) \cdot g(x)$ is continuous at a

[3] $\dfrac{f(x)}{g(x)}$ is continuous at a (if $g(x) \neq 0\ \forall x \in X$)

[4] for any constant $\lambda \in R$ the function $\lambda f(x)$ is continuous at a,

[5] for any natural number $k \in N$ the function $(f(x))^k \equiv f^k(x)$ is continuous at a.

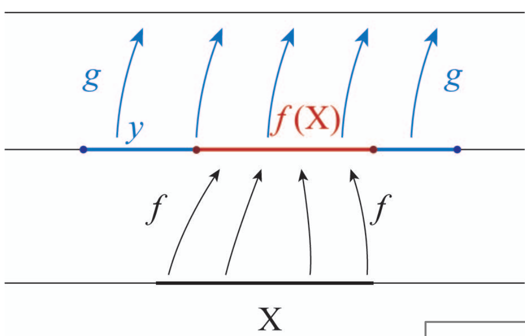
Def. f is defined on X. We say that f is continuous on X if f is continuous at every point $a \in X$.

Obviously, if f continuous on X, then f continuous on any subset $\Omega \subset X$.

Consequence from theorem2. f, g and are continuous on X, then

$f \pm g \,\|\, f \cdot g \,\|\, f/g$ (if $g \neq 0$ on X) $\|\, \lambda f(x)\ (\lambda \in R) \,\|\, f^k(x)\ (k \in N)$ are continuous on X.

Composite function. Some function f is defined on X. let's consider the set of all values $\{f(x) \,\|\, x \in X\} \equiv f(X)$ **[pict6]**. Let Y is any set which contains the set of all the values $f(X)$, and a function $g(y)$ is defined on Y. Then for any number $x \in X$ the number $g(f(x))$ is defined, and we have the function $g(f(x))$ of argument x which is defined on X, this function is called a composite function.



pict.6

Theorem3 [Continuity of a composite function]. $f(x)$ is defined on X and $g(y)$ is defined on $\mathrm{Y} \supset f(\mathrm{X})$. If $f(x)$ is continuous at $a \in \mathrm{X}$ and $g(y)$ is continuous at $f(a)$, then $g(f(x))$ is continuous at $a \in \mathrm{X}$.

Proof. Let's use **the main definition of a continuous function [B]** for $g(f(x))$. We fix an arbitrary sequence $\{x_n\} \subset \mathrm{X} \| \{x_n\} \to a$, if we show that $\{g(f(x_n))\} \to g(f(a))$, then $g(f(x))$ is continuous at a. As $f(x)$ is continuous at a, then from $\{x_n\} \subset \mathrm{X} \| \{x_n\} \to a$ follows that $\{f(x_n)\} \to f(a)$. The sequence $\{f(x_n)\}$ is a sequence of Y which goes to $f(a) \in \mathrm{Y}$, and $g(y)$ is continuous at $f(a)$, then there must be $\{g(f(x_n))\} \to g(f(a))$. Everything is proved.

Consequence1. $f(x)$ is defined on X, and $g(y)$ is defined on $\mathrm{Y} \supset f(\mathrm{X})$.

If $f(x)$ is continuous on X, and $g(y)$ is continuous on Y, then $g(f(x))$ is continuous on X.

Improvement of the Theorem3 [Limit of a composite function].

$\lim_{x \to a \| x \in \mathrm{X}} f(x) = \mathrm{A}$ and $g(y)$ is defined on Y, where $f(\mathrm{X}) \subset \mathrm{Y}$ and $\mathrm{A} \in \mathrm{Y}$.

If $g(y)$ is continuous at A, then $\lim_{x \to a \| x \in \mathrm{X}} g(f(x)) = g(\mathrm{A})$.

Exercise4. Show that the **improvement of the theorem3** is true.

We have provided above the most general and the most important limit-definition:
$\lim_{x \to a \| x \in \mathrm{X}} f(x) = \mathrm{A}$.

Ordinary limit. Let now X is some deleted neighborhood of a, so $\mathrm{X} \equiv D_R(a)$.

The Reader can easily check that the equivalent limit definitions **Def[A]** and **Def[B]** turn into:

Def [ordinary limit A]. If for any sequence $\{x_n\} \in D_R(a) \| \{x_n\} \to a$ the sequence $\{f(x_n)\} \to \mathrm{A}$, then we write $\lim_{x \to a} f(x) = \mathrm{A}$.

Def [ordinary limit B]. If for any (small) $\varepsilon > 0$ there exist $\delta : 0 < \delta < R$ such that for any $x \in D_\delta(a)$ we have $|f(x) - \mathrm{A}| < \varepsilon$, then we write $\lim_{x \to a} f(x) = \mathrm{A}$.

Left and right limits. Here are two similar definitions. Let f is defined on some "left neighborhood" $(a - R, a)$ of a (for the left limit). Or f is defined on some "right neighborhood" $(a, a + R)$ of a (for the right limit).

Def [left limit A]. If for any sequence $\{x_n\} \in (a - R, a) \| \{x_n\} \to a$ the sequence $\{f(x_n)\} \to \mathrm{A}$, then we write $\lim_{x \to a-} f(x) = \mathrm{A}$ and we say that A is a left limit of $f(x)$ at the point a.

[right limit A]. If for any sequence $\{x_n\} \in (a, a + R) \| \{x_n\} \to a$ the sequence $\{f(x_n)\} \to \mathrm{A}$, then we write $\lim_{x \to a+} f(x) = \mathrm{A}$ and we say that A is a right limit of $f(x)$ at the point a.

[left limit B]. If for any (small) $\varepsilon > 0$ there exist $\delta : 0 < \delta < R$ such that for any $x \in (a-\delta, a)$ we have $|f(x) - \mathrm{A}| < \varepsilon$, then $\lim_{x\to a-} f(x) = \mathrm{A}$.

[right limit B]. If for any (small) $\varepsilon > 0$ there exist $\delta : 0 < \delta < R$ such that for any $x \in (a, a+\delta)$ we have $|f(x) - \mathrm{A}| < \varepsilon$, then $\lim_{x\to a+} f(x) = \mathrm{A}$.

Exercise5. Both left and right limits at a exist and both limits are equal to $\mathrm{A} \Leftrightarrow$ The ordinary limit at a exists and it is equal to A. **Shortly**: $\lim_{x\to a-} f(x) = \lim_{x\to a+} f(x) = \mathrm{A} \Leftrightarrow \lim_{x\to a} f(x) = \mathrm{A}$

Ordinary continuity. f is defined on some neighborhood $\mathrm{O}_R(a)$.

Def [A]. If for any sequence $\{x_n\} \in \mathrm{O}_R(a) \,||\, \{x_n\} \to a$ the sequence $\{f(x_n)\} \to f(a)$, then f is continuous at a.

Def [B]. If for any (small) $\varepsilon > 0$ there exist $\delta : 0 < \delta < R$ such that for any $x \in \mathrm{O}_\delta(a)$ we have $|f(x) - f(a)| < \varepsilon$, then f is continuous at a.

The left\right continuity requires from f to be defined on $(a-R, a]$ or on $[a, a+R)$.
And in these cases we say “ f is left-continuous at a “ or “ f is right-continuous at a”.
From the **exercise5** follows: f is left and right continuous at $a \Leftrightarrow f$ is continuous at a.

Def. Let $\{x_n\}$ is some sequence and $n_1 < n_2 < n_3 < n_4 < ....$ is any increasing sequence of natural numbers, then the sequence $x_{n_1}, x_{n_2}, x_{n_3}, x_{n_4} .... = \{x_{n_k}\}$ is called a subsequence of $\{x_n\}$.

Exercise6. $\{x_n\}$ has a limit a, then any it’s subsequence $\{x_{n_k}\} \subset \{x_n\}$ has the same limit a.
The converse assertion is obviously not true, if some subsequence $\{x_{n_k}\} \to a$, the sequence $\{x_n\}$ mustn’t converge to a.

Lemma1. Any bounded sequence has a convergent subsequence.

Proof. Let $\{x_n\}$ is a bounded sequence. Then $|x_n| \le C$ for any $n \Leftrightarrow -C \le x_n \le C$. Let’s divide the segment $[-C, C]$ into two equal parts: $[-C, 0]$ and $[0, C]$. At least one of these segments (for example $[0, C]$) contains infinitely many terms of $\{x_k\}$, there must be the term with the least number n_1, we fix it $x_{n_1} \in [0, C]$ it is the first term of our subsequence.
Let’s divide now the segment $[0, C]$ into two equal parts: $[0, C/2]$ and $[C/2, C]$. At least one of these parts contains infinitely many terms $\{x_k\}$ (if both parts contain infinitely many terms, then we choose the right one), let it be $[C/2, C]$. There must be the term with the least number $n_2 > n_1$, we fix it $x_{n_2} \in [C/2, C]$ and we continue the process. We will get the sequence of nested segments, each next segment is a half of a previous one, therefore there exist exactly one point h, which is common to all these segments. Obviously $h \in [-C, C]$. And we also have the subsequence $\{x_{n_k}\}$, let’s show that $\{x_{n_k}\} \to h$.

Let's fix an arbitrary positive $\varepsilon > 0$ and consider the neighborhood $O_\varepsilon(h)$, as h is a common point of all segments, therefore $O_\varepsilon(h)$ contains all the segments, starting from some number $\bar{k}$ (really, if $O_\varepsilon(h)$ does not fully contain some segment Δ such that $h \in \Delta$, then the length of the segment Δ is not less than $\varepsilon/2 > 0$, it can't be true for all segments, because their lengths go to zero). Then $O_\varepsilon(h)$ contains the segments with numbers $\bar{k}+1$,$\bar{k}+2$,$\bar{k}+3$ .... Every segment with number m contains the term x_{n_m} of our subsequence, then $O_\varepsilon(h)$ contains all terms of $\{x_{n_k}\}$ with numbers $\bar{k}+1$,$\bar{k}+2$,$\bar{k}+3$ .... And ε is an arbitrary small positive number, then $\{x_{n_k}\} \to h$.

Def[U]. f is defined on X, f is called uniformly continuous on X if for any $\varepsilon > 0$ there exist $\delta > 0$ such that $\forall x_1, x_2 \in X : |x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$.

Assertion1. Uniform continuity on X is "stronger" than the ordinary continuity.
If f is uniformly continuous on X, then f is continuous on X, but not conversely.

Proof. Let f is uniformly continuous on X, let's fix an arbitrary $a \in X$.

And let's take $x_2 \equiv a$ in the **Def[U]**. Then for any $\varepsilon > 0$ there exist $\delta > 0$ such that $\forall x_1 \in X : |x_1 - a| < \delta \Rightarrow |f(x_1) - f(a)| < \varepsilon$ it means exactly that f is continuous at a.

Then f is continuous at every point $a \in X$ and f is continuous on X.

Comment. The converse assertion is not true. If f is continuous on X, then f mustn't be uniformly continuous on X.

Lemma2. A sequence $\{x_n\} \subset [a,b]$ converges to some limit c, then $c \in [a,b]$.

Proof. If $c \notin [a,b]$ then there exist some neighborhood $O_\varepsilon(c)$ which doesn't have any common points with $[a,b]$, this neighborhood contains all the terms of $\{x_n\}$, starting from some number, which is impossible, because $\{x_n\} \subset [a,b]$. This contradiction proves that $c \in [a,b]$.

Cantor's theorem. f is continuous on $[a,b]$, then f is uniformly continuous on $[a,b]$.

Proof. Let's assume that there is no uniform continuity, then there exist some "bad" $\bar{\varepsilon} > 0$ for which it's impossible to find any appropriate $\delta > 0$, it means that for any $\delta > 0$ there exist some points x, z such that $|x - z| < \delta$, but anyway $|f(x) - f(z)| \geq \bar{\varepsilon}$. Let's fix an arbitrary positive sequence $\{\delta_n\} \to 0$, for every δ_n there exist $x_n, z_n \in [a,b]$ such that $|x_n - z_n| < \delta_n$ and $|f(x_n) - f(z_n)| \geq \bar{\varepsilon}$ **[Z]**.
All the elements of both sequences $\{x_n\}$ and $\{z_n\}$ belong to $[a,b]$, then both these sequences are bounded, then (lemma1) each of these sequences has a convergent subsequence, let's take just one subsequence $\exists \{x_{n_k}\} \to c$, this subsequence defines an increasing sequence of natural numbers $n_1 < n_2 < n_3 < n_4 <$ .... and this sequence defines the subsequences $\{z_{n_k}\}$ and $\{\delta_{n_k}\}$. As $\{\delta_n\} \to 0$, then the subsequence $\{\delta_{n_k}\} \to 0$. We also have $|x_n - z_n| < \delta_n$, from here follows that $|x_{n_k} - z_{n_k}| < \delta_{n_k} \Leftrightarrow -\delta_{n_k} < x_{n_k} - z_{n_k} < \delta_{n_k} \Leftrightarrow x_{n_k} - \delta_{n_k} < z_{n_k} < x_{n_k} + \delta_{n_k}$, both outer sequences

$\{x_{n_k} - \delta_{n_k}\}$ and $\{x_{n_k} + \delta_{n_k}\}$ converge to c (because $\{x_{n_k}\} \to c$ and $\{\delta_n\} \to 0$), then, according to the squeeze theorem for sequences, $\{z_{n_k}\} \to c$.

So we have $\{x_{n_k}\} \to c$ and $\{z_{n_k}\} \to c$, both these sequences belong to $[a,b]$, then (lemma2) $c \in [a,b]$.

As f is continuous at the point c, there must be $\{f(x_{n_k})\} \to f(c)$ and $\{f(z_{n_k})\} \to f(c)$.

As sequences $\{f(x_{n_k})\}$ and $\{f(z_{n_k})\}$ go to the same limit, their difference $\{f(x_{n_k}) - f(z_{n_k})\}$ is an infinitely small sequence.

But from **[Z]** we have $\left|f(x_{n_k}) - f(z_{n_k})\right| \geq \bar{\varepsilon}$, it means that the absolute value of any term of the sequence $\{f(x_{n_k}) - f(z_{n_k})\}$ is not less than $\bar{\varepsilon}$, then $\bar{\varepsilon}$ neighborhood of 0 does not contain any terms of this sequence. Then $\{f(x_{n_k}) - f(z_{n_k})\}$ is **not** an infinitely small, and we have a contradiction. This contradiction proves our theorem.

Def. A function $f(x)$ is defined on X, $f(x)$ is called bounded on X if there exist some $C > 0$ such that $|f(x)| < C \,\|\, \forall x \in \mathrm{X}$.

1-st Weierstrass theorem. $f(x)$ is continuous on $[a,b]$, then f is bounded on $[a,b]$.

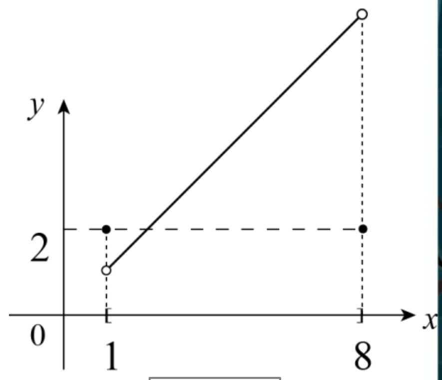
Proof. Let's assume that it is not true, then for any natural number n there exist at least one point $x_n \in [a,b]$ such that $|f(x_n)| > n$. For $n = 1$ there exist $x_1 \in [a,b]$ such that $|f(x_1)| > 1$.

For $n = 2$ there exist $x_2 \in [a,b]$ such that $|f(x_2)| > 2$ and etc. Then we have the sequence $\{x_n\}$ (notice that there we do not require from the terms of this sequence to be different numbers, so there may be for example that $x_1 = x_2$, or $x_2 = x_3$ and etc). All terms of $\{x_n\}$ belong to $[a,b]$ and therefore $\{x_n\}$ is bounded. According to the lemma1, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\} \to h$.

As $\{x_{n_k}\}$ belongs to $[a,b]$ then (lemma2) $h \in [a,b]$. As f is continuous at the point h, from $\{x_{n_k}\} \to h$ follows that $\{f(x_{n_k})\} \to f(h)$. So the sequence $\{f(x_{n_k})\}$ converges, and in the same time we have $|f(x_{n_1})| > n_1$, $|f(x_{n_2})| > n_2$, $|f(x_{n_3})| > n_3$ ..., where n_1, n_2, n_3 ... is an increasing sequence of natural numbers. From here immediately follows that $\{f(x_{n_k})\}$ is **an unbounded** sequence. Then $\{f(x_{n_k})\}$ **does not converge** to any limit (really, any sequence that converges to some number $f(h)$ must be bounded, we have proved it earlier). And we have a contradiction. Everything is proved.

Let some function f is bounded on $[a,b]$. Then the set of all values $\{f(x) \,\|\, x \in [a,b]\}$ is a bounded set, therefore it has the least upper bound M and a greatest lower bound m: so $m \leq f(x) \leq \mathrm{M}$ for any $x \in [a,b]$. But **there is no guarantee** that f reaches the numbers m, M (i.e., there may no be any $x_0 \in [a,b]$ such that $f(x_0) = m$ and there may no be any $\tilde{x}_0 \in [a,b]$ such that that $f(\tilde{x}_0) = \mathrm{M}$).

Example [pict7]. Let's consider the function f on $[1,8]$:
$f(1)=f(8)=2$ and $f(x)\equiv x\ \forall x\in(1,8)$. Here $m=1$ and $\mathrm{M}=8$, but f doesn't reach any of these values on $[1,8]$.



pict.7

There is only a guarantee that $f(x)$ approaches arbitrary close to m (such that $f(x)\ge m$), for some values $x\in[a,b]$. And the closer $f(x)$ goes to m, the less the value $f(x)$ we have.
And similarly, $f(x)$ may not reach the value M on $[a,b]$, but $f(x)$ may approach arbitrary close to M (such that $f(x)\le\mathrm{M}$). And the closer $f(x)$ goes to M, the greater the value $f(x)$ we have.

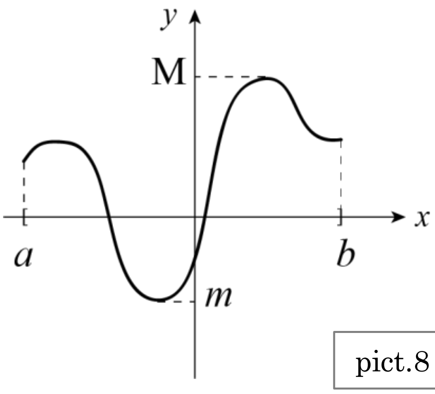
So, for any function f, even if f is bounded on $[a,b]$, the notions " the maximum value of f on $[a,b]$" and " the minimum value of f on $[a,b]$" are not correct. There may no be any maximum/minimum value. But these notions are correct if f is continuous on $[a,b]$ and the 2-nd Weierstrass theorem proves it.

2-nd Weierstrass theorem. $f(x)$ is continuous on $[a,b]$, then f reaches on $[a,b]$ it's maximum M and it's minimum m.

Proof. From the 1-st theorem follows that f is bounded on $[a,b]$. Then the set of all values $\{f(x)\,\|\,x\in[a,b]\}$ is a bounded set, and it must have the greatest lower bound m and the least upper bound M.
[A] Let's show that f reaches M on $[a,b]$. We assume that there is no $x_0\in[a,b]$ such that $f(x_0)=\mathrm{M}$ and we consider the function $g(x)=\dfrac{1}{\mathrm{M}-f(x)}$ on $[a,b]$. The numerator 1 can be considered as a constant function, such function is obviously continuous on $[a,b]$, the denominator $\mathrm{M}-f(x)$ is also a continuous on $[a,b]$ function (because f is continuous on $[a,b]$).

So $g(x)=\dfrac{1}{\mathrm{M}-f(x)}$ is a quotient of two continuous functions on $[a,b]$. The function $\mathrm{M}-f(x)$ is positive everywhere on $[a,b]$, and as $f(x)$ may approach arbitrary close to M, the difference $\mathrm{M}-f(x)$ may be an arbitrary small positive number, then $\dfrac{1}{\mathrm{M}-f(x)}$ may be an arbitrary big positive number. So $g(x)$ is not bounded on $[a,b]$, and in the same time $g(x)$ is continuous on $[a,b]$, it contradicts to the 1-st Weierstrass theorem. This contradiction proves that f actually reaches it's maximum M at some point of $[a,b]$. **[B]** In order to show that f reaches m on $[a,b]$ we can consider the function $\varphi(x)=\dfrac{1}{f(x)-m}$.

Def. Let f is bounded on $[a,b]$, the number $\mathrm{M}-m$, where $\mathrm{M}=\sup_{[a,b]} f(x)$ and $m=\inf_{[a,b]} f(x)$ is called an **oscillation** of f on $[a,b]$. And we denote it: $\omega_{[a,b]}(f)\equiv \mathrm{M}-m$. **[pict8].**

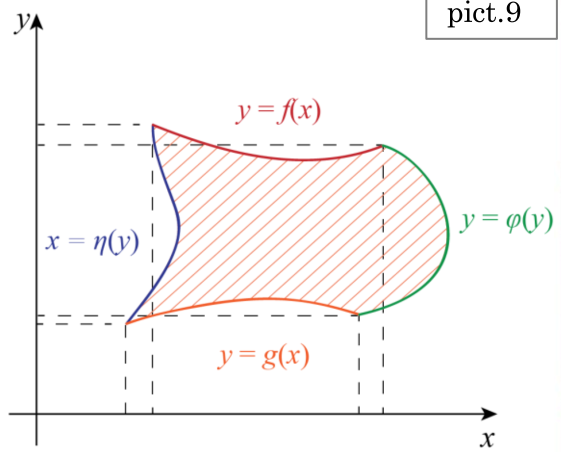


pict.8

Exercise1. f is bounded on $[a,b]$, and for any $x_1,x_2\in[a,b]$ we have $|f(x_1)-f(x_2)|<\mathrm{T}$, where T is some fixed positive number. Then $\sup_{[a,b]} f(x)-\inf_{[a,b]} f(x)\le \mathrm{T} \Leftrightarrow \omega_{[a,b]}(f)\le \mathrm{T}$.

Theorem4. Let Ω is some figure on the plane, the boundary $\partial\Omega$ consists of several graphs of continuous functions like $y=f(x)\,\|\,x\in[a,b]$ or $x=g(y)\,\|\,y\in[c,d]$ **[pict9]**.
Then Ω is measurable.

pict.9

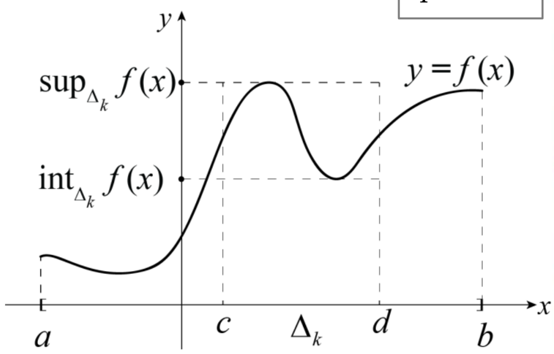


Proof. It's enough to show that $\partial\Omega$ is a zero area figure. We will show that for any positive $\varepsilon>0$ the graph $y=f(x)\,\|\,x\in[a,b]$ can be covered by several rectangles, without common internal points which total area is not greater than ε. All these rectangles together form an external measurable figure which area is not greater than ε. Then the graph of $y=f(x)\,\|\,x\in[a,b]$ is measurable and it's area is zero. Then the boundary $\partial\Omega$ of the initial figure Ω is a union of several zero area figures (several graphs), then the boundary $\partial\Omega$ is also a zero area figure, then Ω is measurable. Let's start.

We have $y=f(x)\,\|\,x\in[a,b]$ is a continuous function, then (**Cantor's theorem**) f is uniformly continuous on $[a,b]$. Let's fix an arbitrary small positive $\varepsilon>0$. For the positive number $\dfrac{\varepsilon}{b-a}>0$ there exist $\delta>0$ such that $\forall x_1,x_2\in[a,b]:|x_1-x_2|<\delta \Rightarrow |f(x_1)-f(x_2)|<\dfrac{\varepsilon}{b-a}$.

Let's divide the initial segment $[a,b]$ into several consecutive segments $\Delta_1,\Delta_2,\Delta_3 \ldots.\ \Delta_m$ which go one after another, and the length of each segment Δ_k is less than δ. Then for any points $x_1,x_2\in\Delta_k$ we "automatically" have $|x_1-x_2|<\delta$ and therefore for any $x_1,x_2\in\Delta_k$ we have $|f(x_1)-f(x_2)|<\dfrac{\varepsilon}{b-a}$. The function f is continuous on every segment Δ_k, then f is bounded on every Δ_k and the numbers $\sup_{\Delta_k} f(x)$ and $\inf_{\Delta_k} f(x)$ are defined (supremum and infinum of f on Δ_k).

pict.10



From $\forall x_1,x_2\in\Delta_k \Rightarrow |f(x_1)-f(x_2)|<\dfrac{\varepsilon}{b-a}$ (**exercise1**)

follows that $\sup_{\Delta_k} f(x) - \inf_{\Delta_k} f(x) \le \dfrac{\varepsilon}{b-a}$ $[\tilde{V}]$. The graph of f on the segment $\Delta_k \equiv [c,d]$ lies inside the rectangle which is formed by the lines $x = c,\ x = d,\ y = \inf_{\Delta_k} f(x),\ y = \sup_{\Delta_k} f(x)$ **[pict10]**. The area of such rectangle is equal to

$$(d-c)\cdot\left(\sup_{\Delta_k} f(x) - \inf_{\Delta_k} f(x)\right) = |\Delta_k|\cdot\left(\sup_{\Delta_k} f(x) - \inf_{\Delta_k} f(x)\right) \le [\tilde{V}] \le |\Delta_k|\cdot\frac{\varepsilon}{b-a}$$

(here $|\Delta_k|$ is the length of Δ_k).

So, all the graph $y = f(x) \,\|\, x \in [a,b]$ is covered by rectangles, every rectangle is built under some segment Δ_k, the total sum of areas: $(Total\ sum\ of\ areas) \le |\Delta_1|\cdot\dfrac{\varepsilon}{b-a} + |\Delta_2|\cdot\dfrac{\varepsilon}{b-a} + \ldots + |\Delta_k|\cdot\dfrac{\varepsilon}{b-a} =$

$= \dfrac{\varepsilon}{b-a}\cdot\left(|\Delta_1| + |\Delta_2| + \ldots + |\Delta_k|\right) = \dfrac{\varepsilon}{(b-a)}\cdot(b-a) = \varepsilon$. Everything is proved.

Def. f is defined on X, f is called monotonically increasing if for any $x_1, x_2 \in X : x_1 < x_2 \Rightarrow f(x_1) \le f(x_2)$. And f is called “strictly increasing” if for any $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$. And f is called monotonically decreasing if for any $x_1 < x_2 \Rightarrow f(x_1) \ge f(x_2)$ and strictly decreasing if for any $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$.

If f is monotonically increasing/decreasing on X then we can say “ f is monotonic on X”.

If f is strictly increasing/decreasing on X, then we can say “ f is strictly monotonic on X”.

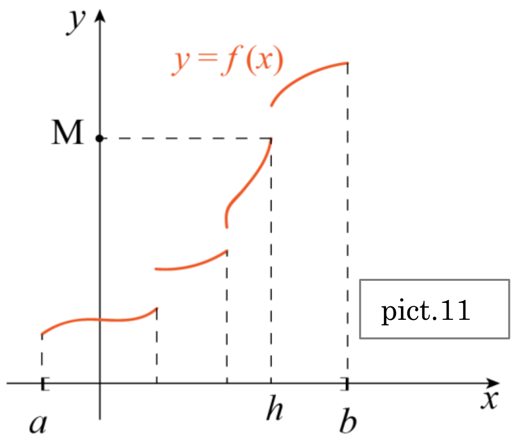
Theorem5: f is monotonically increasing on $[a,b]$. Then at any point $h \in (a,b)$ there exist both left and right limits: $\lim_{x\to h-} f(x)$ and $\lim_{x\to h+} f(x)$. Moreover, the left limit at any point $h \in (a,b)$ is exactly the supremum of all values $\{f(x) \,\|\, x \in [a,h)\}$. And the right limit at any point $h \in (a,b)$ is exactly the infinum of all values $\{f(x) \,\|\, x \in (h,b]\}$.

In particular, at the end points a,b there exist the right limit $\lim_{x\to a+} f(x)$ and the left limit $\lim_{x\to b-} f(x)$.

Proof. We fix an arbitrary $h \in (a,b)$, let’s consider the set of values $\{f(x) \,\|\, x \in [a,h)\}$, this set is bounded above by $f(b)$. Therefore this set has the least upper bound M **[pict11]**.

Let’s show that $\lim_{x\to h-} f(x) = M$. According to the limit-definition, if we show that for any sequence $\{x_n\} \subset [a,h) \,\|\, \{x_n\} \to h$ there must be $\{f(x_n)\} \to M$, then $\lim_{x\to h-} f(x) = M$.

Let’s fix an arbitrary sequence $\{x_n\} \subset [a,h) \,\|\, \{x_n\} \to h$.



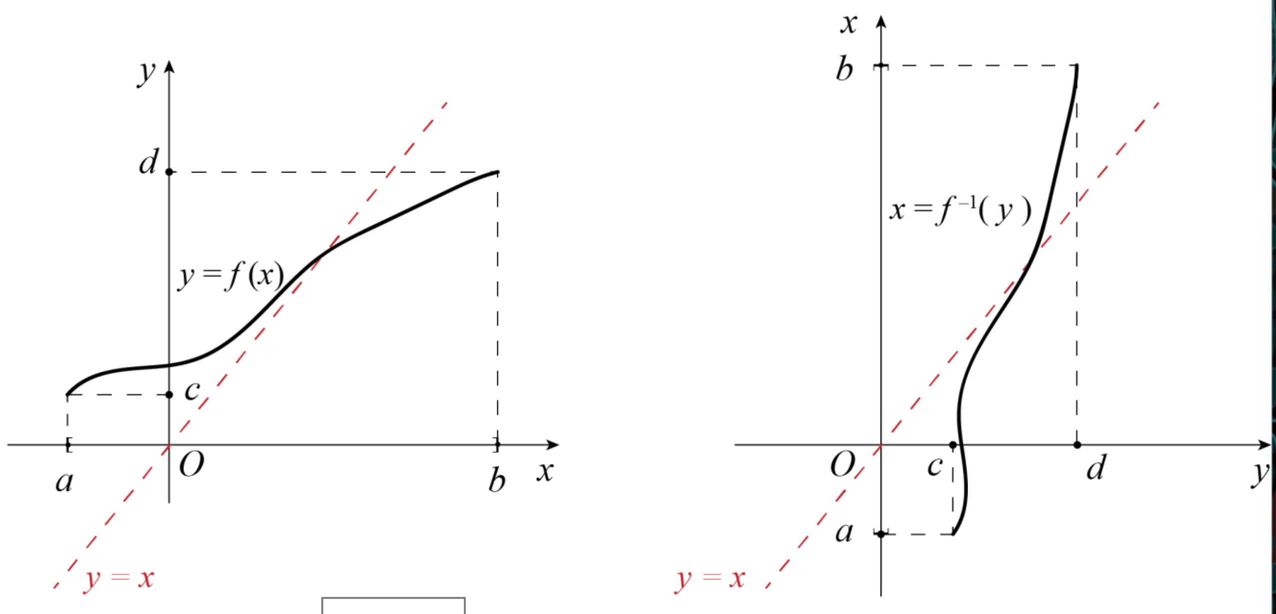
pict.11

Let's fix an arbitrary positive ε, as M is the least upper bound of the set $\{f(x) \| x \in [a,h)\}$, the half interval $(\mathrm{M}-\varepsilon, \mathrm{M}]$ must contain at least one value $f(x_0) \| x_0 \in [a,h)$ (here we use the least upper bound criterion). Let's consider the interval (x_0, h), as $\{x_n\} \to h \| \{x_n\} \subset [a,h)$, the interval (x_0, h) contains all the terms of $\{x_n\}$, starting from some number k. Let's write it:
$\exists k : \forall n > k \Rightarrow x_0 < x_n < h \Rightarrow x_0 < x_n \| x_n \in [a,h)$. As f is monotonically increasing, there must be $f(x_0) \le f(x_n) \| x_n \in [a,h)$, remember that $f(x_0)$ belongs to $(\mathrm{M}-\varepsilon, \mathrm{M}]$, then $f(x_n)$ also belongs to $(\mathrm{M}-\varepsilon, \mathrm{M}]$. Really, we already have $f(x_0) < f(x_n)$ and also $f(x_n) \le \mathrm{M}$, because M is the least upper bound of $\{f(x) \| x \in [a,h)\}$ and $x_n \in [a,h)$. Then $f(x_n)$ belongs to $(\mathrm{M}-\varepsilon, \mathrm{M}]$, and it is true for any $n > k$.
So, for any $n > k$ all the terms of the sequence $\{f(x_n)\}$ belong to $\mathrm{O}_\varepsilon(\mathrm{M}) = (\mathrm{M}-\varepsilon, \mathrm{M}] \cup [\mathrm{M}, \mathrm{M}-\varepsilon)$, and ε is an arbitrary small positive number, then $\{f(x_n)\} \to \mathrm{M}$. We have proved that
$\lim_{x \to h-} f(x) = \mathrm{M}$.
Similarly we can show that there exist the right limit $\lim_{x \to h+0} f(x) = m$. We consider the set $\{f(x) \| x \in (h,b]\}$. This set is bounded below by $f(a)$, then it has the greatest lower bound m.
And we can show that $\lim_{x \to h+} f(x) = m$.

Exercise2. Formulate the theorem, which is similar to the **theorem5**, for a monotonically decreasing function f on $[a,b]$.

Let's remind that if some function $f(x)$ is one-to-one $f : \mathrm{X} \to \mathrm{Y}$, then we can speak about an inverse function $f^{-1} : \mathrm{Y} \to \mathrm{X}$.

Graph of an inverse function. In practice it's very convenient to use the next idea. Suppose we have a graph of one-to-one function f on X and we need to understand quickly how the graph of the inverse function f^{-1} looks. All we need to do is to reflect all our picture over the line $y = x$, such symmetry changes the places of Ox, Oy **[pict12]**, the new graph (after the symmetry) is the graph of f^{-1}.



pict.12

Inverse function theorem. A function $y = f(x)$ is **strictly** increasing and continuous on $[a,b]$. Then: **[A]** f is one-to-one mapping $f:[a,b] \to [f(a), f(b)]$ and

[B] An inverse function $x = f^{-1}(y)$, which is defined on $[f(a), f(b)]$, is also strictly increasing and continuous on $[f(a), f(b)]$.

Proof. Let's fix any number $\mathrm{T} \in [f(a), f(b)]$, according to the (Book1,page145, consequence1 **intermediate values of a continuous function**), there exist $c \in [a,b]$ such that $f(c) = \mathrm{T}$. It means that $f:[a,b] \to [f(a), f(b)]$ covers the segment $[f(a), f(b)]$. And f obviously doesn't "glue together" numbers from $[a,b]$. Really, for any $c \neq d$ from $[a,b]$ we have $f(c) \neq f(d)$
(**If** $c < d \Rightarrow f(c) < f(d)$, **If** $d < c \Rightarrow f(d) < f(c)$, because f is strictly increasing).
Then f is one-to-one mapping $[a,b] \to [f(a), f(b)]$. So **[A]** is proved.

As f is one-to-one, then there exist the inverse one-to-one mapping $f^{-1}:[f(a), f(b)] \to [a,b]$ which we denote as $x = f^{-1}(y)$. Let's show that this function is also strictly increasing.

Let $y_1, y_2 \in [f(a), f(b)] \,\|\, y_1 < y_2$, let's show that $f^{-1}(y_1) < f^{-1}(y_2)$ **[Z]**.

Let's notice that $f^{-1}(y_1)$, $f^{-1}(y_2)$ are some numbers from $[a,b]$. **If** $f^{-1}(y_1) = f^{-1}(y_2)$, then $f\left(f^{-1}(y_1)\right) = f\left(f^{-1}(y_2)\right) \Leftrightarrow y_1 = y_2$, and we have a contradiction. **If** $f^{-1}(y_1) > f^{-1}(y_2)$, then $f\left(f^{-1}(y_1)\right) > f\left(f^{-1}(y_2)\right) \Leftrightarrow y_1 > y_2$, and we have a contradiction again. The last variant is $f^{-1}(y_1) < f^{-1}(y_2)$ (and in this case we don't have any contradiction). So $x = f^{-1}(y)$ is a strictly increasing function.

Let's show that $f^{-1}(y)$ is continuous on $[f(a), f(b)]$. We need to show that at any concrete point $y_0 \in [f(a), f(b)]$ the limit $\lim_{y \to y_0} f^{-1}(y)$ exists and this limit is equal to $f(y_0)$.
(when $y_0 = f(a)$ or $y_0 = f(b)$ we speak about right/left limit).

Let's consider the most general case, we fix any point $y_0 \in (f(a), f(b))$ inside the segment $[f(a), f(b)]$. As f^{-1} is strictly increasing (and therefore monotonically increasing), according to the **theorem4**, there exist both left and right limits $\lim_{y \to y_0-} f^{-1}(y)$ and $\lim_{y \to y_0+} f^{-1}(y)$.

And the most important: $\lim_{y \to y_0-} f^{-1}(y) = \sup\{f^{-1}(y) \,\|\, y \in [f(a), y_0)\}$ **[L1]** and

$\lim_{y \to y_0+} f^{-1}(y) = \inf\{f^{-1}(y) \,\|\, y \in (y_0, f(b)]\}$ **[L2].**

As f^{-1} is strictly increasing on $[f(a), f(b)]$, then any element of the set $\{f^{-1}(y) \,\|\, y \in [f(a), y_0)\}$ is strictly less than $f^{-1}(y_0)$, and $f^{-1}(y_0)$ is strictly less than any element of the set $\{f^{-1}(y) \,\|\, y \in (y_0, f(b)]\}$. From here immediately follows that
$\sup\{f^{-1}(y) \,\|\, y \in [f(a), y_0)\} \leq f^{-1}(y_0) \leq \inf\{f^{-1}(y) \,\|\, y \in (y_0, f(b)]\}$ **[J]**.

If $\sup\{f^{-1}(y) \| y \in [f(a), y_0)\} < f^{-1}(y_0)$, then there exist some number x_1 between these numbers $\sup\{f^{-1}(y) \| y \in [f(a), y_0)\} < x_1 < f^{-1}(y_0) \le \inf\{f^{-1}(y) \| y \in (y_0, f(b)]\}$ **[J1]**.

Then from **[J1]** we see that x_1 is not a value of the function f^{-1} at any point $y \in [f(a), f(b)]$, so f^{-1} doesn't reach the value x_1 at any point of $[f(a), f(b)]$, in the same time x_1 is a point of the segment $[a,b]$, then f^{-1} is not one-to-one $[f(a), f(b)] \to [a,b]$ and we have a contradiction. Then in **[J]** there must be an equality sign "=", i.e., $\sup\{f^{-1}(y) \| y \in [f(a), y_0)\} = f^{-1}(y_0) \le \inf\{f^{-1}(y) \| y \in (y_0, f(b)]\}$. If we assume that the second sign "≤" is actually a sign "<" we will get a similar contradiction. So, the second sign "≤" must be also "=". Now we have $\sup\{f^{-1}(y) \| y \in [f(a), y_0)\} = f^{-1}(y_0) = \inf\{f^{-1}(y) \| y \in (y_0, f(b)]\}$ **[J2]**.

Let's sum up. From **[L1]** and **[L2]** and **[J2]** we have:

$\lim_{y \to y_0-} f^{-1}(y) = f^{-1}(y_0) = \lim_{y \to y_0+} f^{-1}(y)$ **[V].** As both left and right limits at y_0 exist, and these limits are equal to $f^{-1}(y_0)$, then the ordinary limit at y_0 exists, and it is equal to $f^{-1}(y_0)$.

So $\lim_{y \to y_0} f^{-1}(y) = f^{-1}(y_0)$, it means exactly that $f^{-1}(y)$ is continuous at y_0.

So $f^{-1}(y)$ is continuous at any point $y_0 \in (f(a), f(b))$. Similarly f^{-1} is right-continuous at $f(a)$ (the proof is similar and even more simple), and f^{-1} is left-continuous at $f(b)$.

Then f^{-1} is continuous on $[f(a), f(b)]$.

Exercise3. Formulate the theorem, which is similar to the **Inverse function theorem**, for a strictly decreasing function f on $[a,b]$.

Exercise4. f is continuous on $[a,b]$. Show that f is strictly monotonic on $[a,b]$ if it reaches different values at different points, i.e., for any $x_1 \ne x_2$ from $[a,b]$ we have $f(x_1) \ne f(x_2)$.

Limits and equivalent functions

Def. We say that $\{x_n\}$ goes to infinity ∞ if for any (big) $C > 0$ there exist some number k, starting from which $\forall n > k$ we have $|x_n| > C$. And we write: $\{x_n\} \to \infty$ or $\lim_{n\to\infty} x_n = \infty$.

In particular, we say that $\{x_n\}$ goes to plus infinity $+\infty$ if for any $\forall C > 0\ \exists k : \forall n > k \Rightarrow\ x_n > C$ and we write $\{x_n\} \to +\infty$ or $\lim_{n\to\infty} x_n = +\infty$. And we say that $\{x_n\}$ goes to minus infinity $-\infty$ if for any $\forall C > 0\ \exists k : \forall n > k \Rightarrow\ x_n < -C$ and we write $\{x_n\} \to -\infty$ or $\lim_{n\to\infty} x_n = -\infty$.

Notice! When we say "$\{x_n\}$ converges" or "$\{x_n\}$ has a limit" we imply that there exist some concrete real number a such that $\lim x_n = a$. The symbols $\infty/+\infty/-\infty$ are **not** limits!
These symbols are just auxiliary symbols, and we use them in order to describe the behavior of our sequence $\{x_n\}$. And when we say that some sequence converges (has a limit) we always imply that it is a "finite limit", which is some concrete real number a.

It's easy to notice that $+\infty$ and $-\infty$ are more detail cases of ∞. So when we have $\{x_n\} \to +\infty$ we can also write $\{x_n\} \to \infty$, and when we have $\{x_n\} \to -\infty$ we can also write $\{x_n\} \to \infty$.

Def. When some sequence goes to ∞ (in particular to $+\infty$ or $-\infty$) we say that we have an "infinitely large sequence".

There exist the similar definition for functions:

Def. f is defined on X and a is a limit point of X. If for any sequence $\{x_n\} \in \mathrm{X} \,\|\, x_n \neq a\ \forall n \,\|\, \{x_n\} \to a$ we have $\{f(x_n)\} \to \infty/+\infty/-\infty$, then we say that $f(x)$ goes to infinity/plus infinity/minus infinity when x goes to a, and we write $\lim_{x\to a\|x\in \mathrm{X}} f(x) = \infty/+\infty/-\infty$.

And again, when we say "f has a limit at a", we always imply that f has some "finite limit" A, which is a **concrete** real number. The symbols $\infty/+\infty/-\infty$ are not limits, these are just auxiliary symbols which we use to describe the behavior of our function f when x approaches to a (by staying in X).

And the equivalent definition: $\lim_{x\to a\|x\in \mathrm{X}} f(x) = \infty/+\infty/-\infty$ if for any (big) number $C > 0$ there exist some $\delta > 0$ such that $\forall x \in \mathrm{X} : x \in D_\delta(a)$ we have $|f(x)| > C //\ \ f(x) > C\ //\ f(x) < -C$.

And there are standard variations: if, for example, $f(x)$ is defined on some $(a-R, a)$ and $f(x)$ goes to plus infinity when x goes to a from the left, then we write $\lim_{x\to a-} f(x) = +\infty$.
If $f(x)$ is defined in some $(a-R, a+R)$ and $f(x)$ goes to minus infinity when x goes to a (from any side), then we write $\lim_{x\to a} f(x) = -\infty$.

Let's consider the function $f(x)=1/x$, this function is defined on $(-\infty,0)\cup(0,+\infty)$ and when x goes to zero from the right, the value $1/x$ goes to plus infinity, then we can write $\lim_{x\to 0+}(1/x)=+\infty$. When x goes to zero from the left, the value $1/x$ goes to minus infinity, then $\lim_{x\to 0-}(1/x)=-\infty$.

And finally, let $f(x)$ is defined on some set $(R,+\infty)$. If there exist some number A such that: For any (small) $\varepsilon>0$ there exist $C>R$ such that $\forall x\in(C,+\infty)\Rightarrow|f(x)-\mathrm{A}|<\varepsilon$, then we say " $f(x)$ goes to A when x goes to $+\infty$ " and we write $\lim_{x\to+\infty} f(x)=\mathrm{A}$.

Let $f(x)$ is defined on $(-\infty,-R)$ (here $-R<0$). If there exist A such that $\forall\varepsilon>0\ \exists -C<-R:\forall x\in(-\infty,-C)\Rightarrow|f(x)-\mathrm{A}|<\varepsilon$, then we say " $f(x)$ goes to A when x goes to $-\infty$ " and we write $\lim_{x\to-\infty} f(x)=\mathrm{A}$.

Let $f(x)$ is defined on $(-\infty,-R)\cup(R,+\infty)$. If there exist A such that $\forall\varepsilon>0\ \exists C>R:\forall x\in(-\infty,-C)\cup(C,+\infty)\Rightarrow|f(x)-\mathrm{A}|<\varepsilon$, then we say " $f(x)$ goes to A when x goes to ∞ " and we write $\lim_{x\to\infty} f(x)=\mathrm{A}$.

Exercise1. By using the previous information, define next the limits: $\lim_{x\to\infty/+\infty/-\infty} f(x)=\infty/+\infty/-\infty$.

For practical purposes it's very important to define infinitely small functions.

Def. $\alpha(x)$ is defined on X and a is a limit point of X. If $\lim_{x\to a\|x\in \mathrm{X}}\alpha(x)=0$, then we say that $\alpha(x)$ is "infinitely small when $x\to a$" ("infinitely small when x goes to a").

Exercise2. Let $\lim_{x\to a\|x\in\mathrm{X}} f(x)=\mathrm{A}$, then the function $\alpha(x)\equiv f(x)-\mathrm{A}\ \forall x\in\mathrm{X}$ is infinitely small when $x\to a$. So, for any function $f(x)$ such that $\lim_{x\to a\|x\in\mathrm{X}} f(x)=\mathrm{A}$ the next representation on X is possible: $f(x)=\mathrm{A}+\alpha(x)$, where $\alpha(x)$ is defined on X and $\alpha(x)$ is infinitely small when $x\to a$.

And conversely. If $f(x)=\mathrm{A}+\alpha(x)$, where $\alpha(x)$ is infinitely small when $x\to a$, then $\lim_{x\to a\|x\in\mathrm{X}} f(x)=\mathrm{A}$. It immediately follows from simpliest properties of limits, really: $\lim_{x\to a\|x\in\mathrm{X}} f(x)=\lim_{x\to a\|x\in\mathrm{X}}(\mathrm{A}+\alpha(x))=\lim_{x\to a\|x\in\mathrm{X}}\mathrm{A}+\lim_{x\to a\|x\in\mathrm{X}}\alpha(x)=\mathrm{A}+0=\mathrm{A}$.

Exercise3. If $\alpha(x)$ and $\beta(x)$ are infinitely small when $x\to a$, then their sum and product $\alpha(x)+\beta(x)$, $\alpha(x)\cdot\beta(x)$ are also infinitely small when $x\to a$.

Exercise4. $\alpha(x)\|x\in\mathrm{X}$ is infinitely small when $x\to a$. And $g(x)$ is bounded on X, then $\alpha(x)\cdot g(x)$ is infinitely small when $x\to a$.

From the **exercise4** follows that: if $\alpha(x)$ is infinitely when $x\to a$, then for any constant $\lambda\in R$ the function $\lambda\alpha(x)$ is also infinitely small when $x\to a$.

Local comparison of functions

Let $f(x), g(x)$ are defined on X and $g(x) \neq 0 \,||\, \forall x \in \mathrm{X}$, and a is a limit point of X.

Def. Suppose that the limit $\lim_{x\to a || x\in \mathrm{X}} \dfrac{f(x)}{g(x)} = k$ exists (it means that k is some concrete real number). **If** $k \neq 0$, then we say that f and g **have the same order** when $x \to a$ ("when x goes to a") and we can use the symbol $f = O(g)$.
In particular, when $k = 1$ we say that f and g are **equivalent** when $x \to a$, and we can write $f \approx g$. And finally, if $k = 0$, we say that f is **negligible with respect to** g when $x \to a$, and we can use the symbol $f = o(g)$.

This simple definition has a great importance in math and it has a very simple meaning.

Let f is **negligible with respect to** g, so $f = o(g) \Leftrightarrow \lim_{x\to a||x\in \mathrm{X}} \dfrac{f(x)}{g(x)} = 0$, then $\dfrac{f(x)}{g(x)}$ is infinitely small when $x \to a$, so the values of $\dfrac{f(x)}{g(x)}$ near the point a approach arbitrary close to zero, it means that near the point a, at any concrete point $x \in \mathrm{X}$, the value $f(x)$ is so much less (in absolute value) than the value $g(x)$, that their ratio is almost zero. And even more, for any small $\varepsilon > 0$ there exist some deleted neighborhood of a such that for any point $x \in \mathrm{X}$ from this neighborhood, the absolute value of $f(x)/g(x)$ is less than ε.

Let f is **negligible with respect to** g when $x \to a$. So we have $\lim_{x\to a||x\in \mathrm{X}} \dfrac{f(x)}{g(x)} = 0$, it means that the ratio $\dfrac{f(x)}{g(x)} \equiv \alpha(x)$ is infinitely small when $x \to a$, from here $f(x) = \alpha(x)\cdot g(x)$ - and we can use this representation instead of $f(x)$ in our reasonings, so instead of $f(x)$ we will write $\alpha(x)\cdot g(x)$, where $\alpha(x)$ is infinitely small when $x \to a$. In many cases this approach is very convenient.

And conversely. If we have the representation $f(x) = \alpha(x)\cdot g(x) \,||\, (g(x) \neq 0 \;\; \forall x \in \mathrm{X})$ on X, where $\alpha(x)$ is infinitely small when $x \to a$, then
$\lim_{x\to a||x\in \mathrm{X}} \dfrac{f(x)}{g(x)} = \lim_{x\to a||x\in \mathrm{X}} \dfrac{\alpha(x)\cdot g(x)}{g(x)} = \lim_{x\to a||x\in \mathrm{X}} \alpha(x) = 0$ and f is **negligible with respect to** g when $x \to a$.

Exercise5. Explain the meaning of the other symbols "O" and "$\approx$" in the same manner as we explained above the meaning of "o".

The next theorem for equivalent functions may be very helpful in many cases:

Theorem1. $f(x), f_1(x)$ and $g(x), g_1(x)$ are defined on X and a is a limit point of X and $g(x) \neq 0,\ g_1(x) \neq 0\ \forall x \in \mathrm{X}$. And $f(x) \approx f_1(x)$ and $g(x) \approx g_1(x)$ when $x \to a$.

If the limit $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{f_1(x)}{g_1(x)}$ exists, then $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{f(x)}{g(x)}$ also exists and these limits are equal.

Consequence. When we calculate some limit $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{f(x)}{g(x)}$ we can replace functions f, g by any equivalent functions f_1, g_1 (so, $f \approx f_1$ and $g \approx g_1$ when $x \to a$) and calculate the new limit, if it exists, then the initial limit also exists and their values are equal.

Proof. We have $f(x) \approx f_1(x)$ and $g(x) \approx g_1(x)$ when $x \to a$. Let's designate $\dfrac{f(x)}{f_1(x)} \equiv h(x)$ and $\dfrac{g(x)}{g_1(x)} \equiv v(x)$, we know that $\lim_{x \to a \| x \in \mathrm{X}} h(x) = 1,\ \lim_{x \to a \| x \in \mathrm{X}} v(x) = 1$ and obviously $v(x) \neq 0$ on X.

The ratio $\dfrac{h(x)}{v(x)}$ is defined everywhere on X and $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{h(x)}{v(x)} = \dfrac{\lim_{x \to a \| x \in \mathrm{X}} h(x)}{\lim_{x \to a \| x \in \mathrm{X}} v(x)} = \dfrac{1}{1} = 1$ **[T]**.

Let the limit $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{f_1(x)}{g_1(x)}$ exists, so $\lim_{x \to a \| x \in \mathrm{X}} \dfrac{f_1(x)}{g_1(x)} = \mathrm{A}$. By using **[T]** and the simpliest properties of limits we can write:

$$\lim_{x \to a \| x \in \mathrm{X}} \frac{f(x)}{g(x)} = \lim_{x \to a \| x \in \mathrm{X}} \left(\frac{h(x)}{v(x)}\right) \cdot \left(\frac{f_1(x)}{g_1(x)}\right) = \lim_{x \to a \| x \in \mathrm{X}} \left(\frac{h(x)}{v(x)}\right) \cdot \lim_{x \to a \| x \in \mathrm{X}} \left(\frac{f_1(x)}{g_1(x)}\right) = 1 \cdot \mathrm{A} = \mathrm{A},$$

so the initial limit exists and it is equal to A.

Next, there are similar definitions of symbols "$O, \approx, o$" in the cases when $x \to +\infty$ or $x \to -\infty$, and they also have a great practical value.

Let $f(x), g(x)$ are defined on $(R, +\infty)$ and $g(x) \neq 0 \| \forall x \in (R, +\infty)$.

Def. Suppose that the limit $\lim_{x \to +\infty} \dfrac{f(x)}{g(x)} = k$ exists.

If $k \neq 0$, then we say that f and g **have the same order** when x goes to $+\infty$ and we can use the symbol $f = O(g)$. In particular, when $k = 1$, we say that f and g are **equivalent** when x goes to $+\infty$, and we can write $f \approx g$. And finally, if $k = 0$, we say that f is **negligible with respect** g when x goes to $+\infty$, and we can use the symbol $f = o(g)$.

Exercise6. Formulate the theorem, which is similar to the theorem1, for the limit $\lim_{x \to +\infty} \dfrac{f(x)}{g(x)}$.

Exercise7. Give the definitions of symbols "$O, \approx, o$", when $x \to -\infty$.

Radians

Assertion1. For any $k \in \mathrm{N}$ the function $f(x) = \sqrt[k]{x}$ is continuous on $[0, +\infty)$.

Proof. Let's fix an arbitrary $a \in [0, +\infty)$ and take any sequence $\{x_n\} \in [0, +\infty) \,||\, x_n \neq a \,||\, \{x_n\} \to a$, we need to show that $\sqrt[k]{x_n} \to \sqrt[k]{a}$, then (by definition) $\sqrt[k]{x}$ is continuous at a.

Auxiliary1. If some positive sequence $\{y_n\}$ is infinitely small, then $\{\sqrt[k]{y_n}\}$ is also infinitely small.

Proof. Let's fix an arbitrary positive ε, and take the positive number ε^k, as $\{y_n\}$ is infinitely small, there exist the number m, starting from which $|y_n - 0| < \varepsilon^k \Leftrightarrow y_n < \varepsilon^k \Rightarrow \sqrt[k]{y_n} < \varepsilon \Leftrightarrow \left|\sqrt[k]{y_n} - 0\right| < \varepsilon$. So for any positive ε there exist the number m, starting from which $\left|\sqrt[k]{y_n} - 0\right| < \varepsilon$, then $\{\sqrt[k]{y_n}\}$ is infinitely small.

Auxiliary2. Auxiliary inequality: $\left|\sqrt[k]{a} - \sqrt[k]{b}\right| \leq \sqrt[k]{|a-b|} \quad \forall a \geq 0, \forall b \geq 0$.

Proof. We can assume that $a \geq b$, if not, we can just permute a, b in both (left and right sides). Then we can discard the absolute value $|\ldots|$ signs. And we need to prove that $\sqrt[k]{a} - \sqrt[k]{b} \leq \sqrt[k]{a-b}$, this inequality is equivalent to $\sqrt[k]{a} \leq \sqrt[k]{a-b} + \sqrt[k]{b}$ and this one can be checked by raising both sides to the $k-th$ power.

Let's prove now the assertion1. $\{x_n\} \in [0, +\infty) \,||\, x_n \neq a \,||\, \{x_n\} \to a$, then $|a - x_n|$ is an infinitely small positive sequence, then (auxiliary1) the sequence $\sqrt[k]{|a - x_n|}$ is infinitely small. Then (auxiliary2) $0 \leq \left|\sqrt[k]{a} - \sqrt[k]{x_n}\right| \leq \sqrt[k]{|a - x_n|}$ from the **squeeze theorem for sequences** immediately follows that $\left|\sqrt[k]{a} - \sqrt[k]{x_n}\right|$ is infinitely small. Obviously, for any sequence $\{y_n\}$ the next is true: $\{|y_n|\}$ is infinitely small $\Leftrightarrow \{y_n\}$ is infinitely small. As $\left|\sqrt[k]{a} - \sqrt[k]{x_n}\right|$ is infinitely small $\Rightarrow \sqrt[k]{a} - \sqrt[k]{x_n}$ is infinitely small, then $\left\{\sqrt[k]{x_n}\right\} \to \sqrt[k]{a}$, everything is proved.

In particular, the square root function $f(x) = \sqrt{x}$ is continuous on $[0, +\infty)$.

Assertion2. $\sin\alpha$ is strictly increasing on $(-90^o, 90^o)$.

Proof. Let's fix arbitrary $\alpha_1 < \alpha_2$ from the interval $(-90^o, 90^o)$, let's show that $\sin\alpha_1 < \sin\alpha_2$. This inequality is equivalent to $\sin\alpha_2 - \sin\alpha_1 > 0$. Let's use the formula for difference of sines:

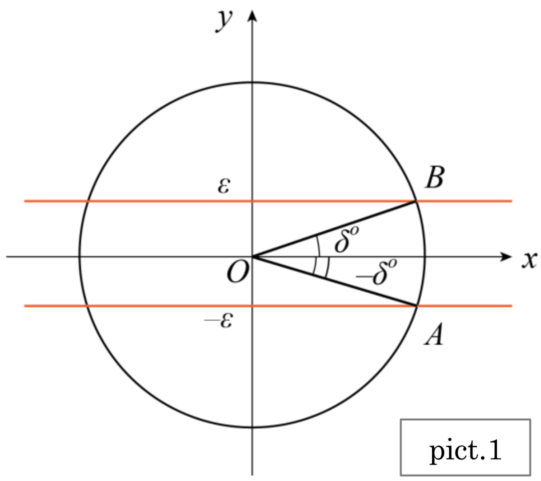
$\sin\alpha_2 - \sin\alpha_1 = 2\cos\dfrac{\alpha_2 + \alpha_1}{2}\sin\dfrac{\alpha_2 - \alpha_1}{2}$ **[T]**. All the factors on the right side of **[T]** are positive.

Really, the angle $\dfrac{\alpha_2+\alpha_1}{2}\in(-90^o,90^o)$, then $\cos\dfrac{\alpha_2+\alpha_1}{2}>0$.

And as $\alpha_2>\alpha_1\Rightarrow\alpha_2-\alpha_1>0^o\Rightarrow\alpha_2-\alpha_1\in(0^o,90^o)$, then $\sin\dfrac{\alpha_2+\alpha_1}{2}>0$, so, the right part of **[T]** is positive. Then $\sin\alpha_2-\sin\alpha_1>0$ and everything is proved.

Assertion3. $\sin\alpha$ is continuous at the point $\tilde{\alpha}=0^o$.

Proof. We want to show that $\lim_{\alpha\to 0^o}\sin\alpha=\sin 0^o=0$.



pict.1

Let's fix an arbitrary positive $\varepsilon\in(0,1)$. We can draw the line $y=\varepsilon$, it intersects the unit circle at two points **[pict1]**, one of these points defines the angle $\angle AOB\in(0^o,90^o)$.

Let's designate $\delta^o\equiv\angle AOB$ and we also consider angle $-\angle AOB=-\delta^o$, which is symmetrical to $\angle AOB$ with respect to Ox. Let's take any angle $\alpha\in(-\delta^o,\delta^o)\Leftrightarrow-\delta^o<\alpha<\delta^o$, as $\sin\alpha$ is strictly increasing on $(-90^o,90^o)$ (lemma2), then

$\sin(-\delta^o)<\sin\alpha<\sin\delta^o\Leftrightarrow-\sin\delta^o<\sin\alpha<\sin\delta^o$ and, according to our construction, $\sin\delta^o=\varepsilon$. So we have $-\varepsilon<\sin\alpha<\varepsilon$.

Let's sum up: for any positive $\varepsilon\in(0,1)$ there exist the positive δ^o such that $\forall\alpha\in(-\delta^o,\delta^o)$ we have $-\varepsilon<\sin\alpha<\varepsilon$. We can rewrite: $\forall\varepsilon\in(0,1)\ \exists\delta>0\ \forall\alpha\in O_\delta(0^o)\Rightarrow|\sin\alpha-0|<\varepsilon$ - it means exactly that $\lim_{\alpha\to 0^o}\sin\alpha=0$.

Assertion4. $\sin\alpha$ is continuous everywhere $(-\infty^o,+\infty^o)$.

Proof. Let's fix an arbitrary angle $\tilde{\alpha}\in(-\infty^o,+\infty^o)$, the value $\sin\tilde{\alpha}$ is a constant. If we show that the function $\sin\alpha-\sin\tilde{\alpha}$ is infinitely small when $\alpha\to\tilde{\alpha}$, then $\lim_{\alpha\to\tilde{\alpha}}\sin\alpha=\sin\tilde{\alpha}$ and $\sin\alpha$ is continuous at $\tilde{\alpha}$. Let's consider $\sin\alpha-\sin\tilde{\alpha}$, we have: $\sin\alpha-\sin\tilde{\alpha}=2\cos\dfrac{\alpha+\tilde{\alpha}}{2}\sin\dfrac{\alpha-\tilde{\alpha}}{2}$ **[M]**.

Let's show that $\sin\dfrac{\alpha-\tilde{\alpha}}{2}$ is infinitely small when $\alpha\to\tilde{\alpha}$. We consider $\sin\dfrac{\alpha-\tilde{\alpha}}{2}$ as a composite function $\mu(\varphi(\alpha))$, here $\varphi(\alpha)\equiv\dfrac{\alpha-\tilde{\alpha}}{2}$ and $\mu(\beta)\equiv\sin\beta$. The function $\varphi(\alpha)$ is continuous at $\tilde{\alpha}$ and $\varphi(\tilde{\alpha})=0^o$ and $\sin\beta$ is continuous at $\beta=0^o$ (assertion3), then the composite function $\mu(\varphi(\alpha))$

is continuous at $\tilde{\alpha}$, and we have $\lim_{\alpha\to\tilde{\alpha}} \mu(\varphi(\alpha)) = \mu(\varphi(\tilde{\alpha})) \Leftrightarrow \lim_{\alpha\to\tilde{\alpha}}\left(\sin\frac{\alpha-\tilde{\alpha}}{2}\right) = \sin\frac{\tilde{\alpha}-\tilde{\alpha}}{2} = 0$

and $\sin\frac{\alpha-\tilde{\alpha}}{2}$ is infinitely small when $\alpha \to \tilde{\alpha}$.

The expression **[M]** can be considered as a product of a bounded function:

$2\cos\frac{\alpha+\tilde{\alpha}}{2}$, $\left|2\cos\frac{\alpha+\tilde{\alpha}}{2}\right| \le 2 \ \forall \alpha \in (-\infty^o, +\infty^o)$ and an infinitely small (when $\alpha \to \tilde{\alpha}$) function

$\sin\frac{\alpha-\tilde{\alpha}}{2}$. Then **[M]** is infinitely small when $\alpha \to \tilde{\alpha}$ function. Everything is proved.

Exericise1. Show that $\cos\alpha$ is continuous everywhere $(-\infty^o, +\infty^o)$.

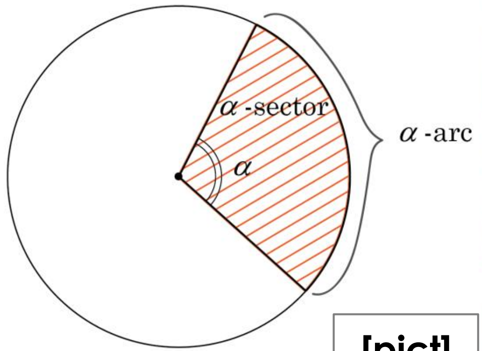
Exericise2. Show that each function $tg\alpha$, $ctg\alpha$ is continuous on it's domain.

The unit circle Ω is a measurable figure. Really, it's boundary consists of two graphs of functions $f(x) = \sqrt{1-x^2} \ || \ x \in [-1,1]$ and $\tilde{f}(x) = -\sqrt{1-x^2} \ || \ x \in [-1,1]$.

Let's show that both these functions are continuous on $[-1,1]$, then (**theorem4**) Ω is measurable.

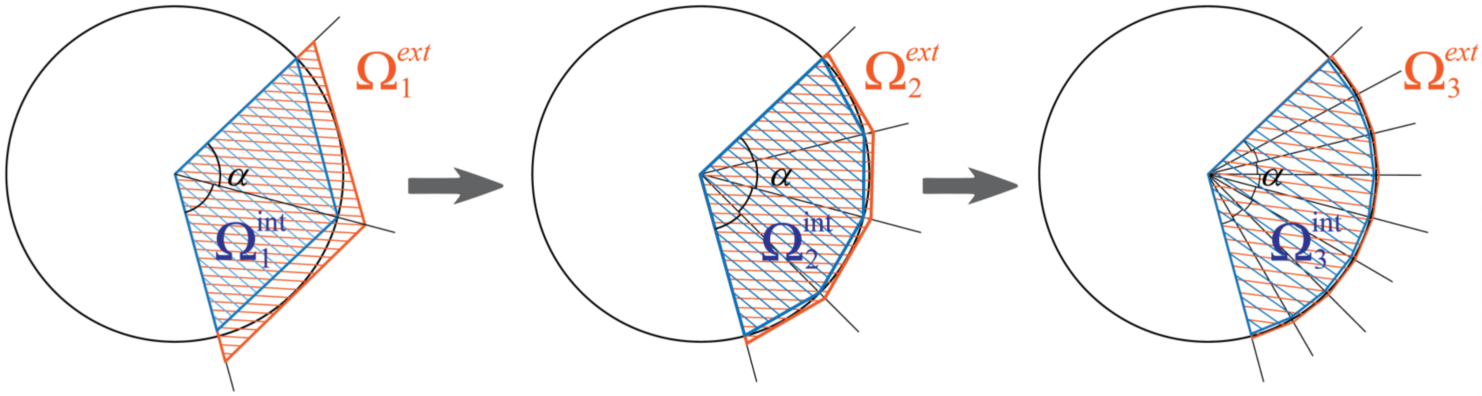
The function x^2 is continuous on $[-1,1]$ (we showed in the $1-st$ book that $x^k \ || \ k \in \mathrm{N}$ is continuous everywhere). Then $\varphi(x) \equiv 1-x^2$ is also continuous on $[-1,1]$, the values of $\varphi(x)$ on $[-1,1]$ belong to the segment $[0,1]$, and the function $\mu(y) \equiv \sqrt{y}$ is continuous on $[0,1]$ (because $\mu(y) \equiv \sqrt{y}$ is continuous on $[0,+\infty)$). Then the composite function $\mu(\varphi(x)) \equiv \sqrt{1-x^2}$ is continuous on $[-1,1]$. And therefore $\tilde{f}(x) = -\sqrt{1-x^2}$ is also continuous on $[-1,1]$.

Def. Ω is a unit circle **[pict]**. Let's fix an arbitrary angle $\alpha \in [0^o, 360^o]$ and let's draw any central angle α of Ω, it defines an arc of a circle, which we will call an α-arc. There also appears a figure, which is called an α-sector.



[pict]

The unit circle Ω is measurable, then it's boundary $\partial\Omega$ is a zero area figure (Book1, "Area construction", **[3-rd criterion of measurability]**). Any α-arc is some part of a zero area figure $\partial\Omega$, then (assertion8, "Area construction") any α-arc is also a zero area figure. Let's consider now any α-sector. It's boundary consists of two equal segments (which are both zero area figures) and an α-arc, then the boundary of any α-sector is a zero area figure, then any α-sector is measurable, and it's area (the number) $S(\alpha)$ is defined. Let's describe the process which gives us the area of any α-sector.



pict.2

The process. Let's fix any α-sector.

For any natural number, which is a power of two 2^n, we divide the angle $\angle\alpha$ into 2^n equal angles. We build the internal figure Ω_n^{int} **[pict2]** by connecting several points on the circle, and the external figure Ω_n^{ext}, by building a tangent line at the middle of every small arc.

The area of any triangle from Ω_n^{int} is equal to $S = \frac{1}{2}\cdot 1\cdot 1\cdot \sin\left(\frac{\alpha}{2^n}\right) = \frac{\sin\left(\alpha/2^n\right)}{2}$. There are 2^n such triangles, so the area $S(\Omega_n^{\text{int}}) = (2^n/2)\cdot\sin(\alpha/2^n) = 2^{n-1}\cdot\sin(\alpha/2^n)$.

Next, any triangle in Ω_n^{ext} has a height 1, it divides the triangle into two equal right triangles, with angles $\frac{\alpha}{2\cdot 2^n}$, so their cathetuses, which are opposite to these angles, are equal to $tg\,\frac{\alpha}{2^{n+1}}$.

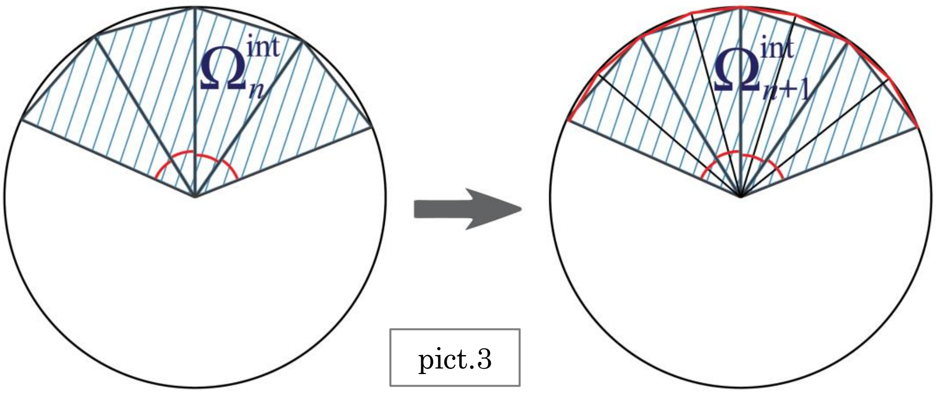
Then the area of each triangle is $\left(tg\,\frac{\alpha}{2^{n+1}} + tg\,\frac{\alpha}{2^{n+1}}\right)\cdot 1\big/2 = tg\,\frac{\alpha}{2^{n+1}}$, and there are 2^n such triangles, so $S(\Omega_n^{ext}) = 2^n\cdot tg(\alpha/2^{n+1})$.

We have two sequences $\{S(\Omega_n^{\text{int}})\} = \{2^{n-1}\cdot\sin(\alpha/2^n)\}$ and $\{S(\Omega_n^{ext})\} = \{2^n\cdot tg(\alpha/2^{n+1})\}$ let's call it an internal and an external sequence.

Assertion5. Both internal and external sequences go to the same limit.

Proof. Let's notice that the internal sequence $\{S(\Omega_n^{\text{int}})\} = \{2^{n-1} \cdot \sin(\alpha/2^n)\}$ is strictly increasing.

Really, for any n we have $\Omega_n^{\text{int}} \subset \Omega_{n+1}^{\text{int}} \Rightarrow S(\Omega_n^{\text{int}}) < S(\Omega_{n+1}^{\text{int}})$ **[pict3]**. Let's explain it.



pict.3

Let we have a figure Ω_n^{int}, now the angle $\angle\alpha$ is divided into 2^n equal central angles.

For each of these angles the angle bisector can be drawn. Then α will be divided into $2 \cdot 2^n = 2^{n+1}$ equal central angles. If we connect now several points on the circle, we will get the figure $\Omega_{n+1}^{\text{int}}$ and obviously $\Omega_n^{\text{int}} \subset \Omega_{n+1}^{\text{int}} \Rightarrow S(\Omega_n^{\text{int}}) < S(\Omega_{n+1}^{\text{int}})$. Notice that the area of $\Omega_{n+1}^{\text{int}}$ differs from the area of Ω_n^{int} by the area of several triangles, which are built on the chords of the figure Ω_n^{int}, that's why the area of $\Omega_{n+1}^{\text{int}}$ is strictly greater than the area of Ω_n^{int}. But in fact, for us it's even enough the estimation $S(\Omega_n^{\text{int}}) \le S(\Omega_{n+1}^{\text{int}})$, and this inequality immediately follows from $\Omega_n^{\text{int}} \subset \Omega_{n+1}^{\text{int}}$.

An internal sequence is also bounded above, really as any internal figure Ω_n^{int} belongs to the α-sector, then it's area is not greater than the area of α-sector, so $S(\Omega_n^{\text{int}}) \le S(\alpha)$.

According to the theorem about a limit of a monotonic sequence, the internal sequence $\{S(\Omega_n^{\text{int}})\} = \{2^{n-1} \cdot \sin(\alpha/2^n)\}$ converges to some limit A. Let's consider now an external sequence $\{S(\Omega_n^{ext})\} = \{2^n \cdot tg(\alpha/2^{n+1})\}$.

Let's take the term of this sequence with number $n-1$, so

$S(\Omega_{n-1}^{ext}) = 2^{n-1} \cdot tg(\alpha/2^n) = 2^{n-1} \cdot \dfrac{\sin(\alpha/2^n)}{\cos(\alpha/2^n)} = \dfrac{(2^{n-1}\sin(\alpha/2^n))}{\cos(\alpha/2^n)} = \dfrac{S(\Omega_n^{\text{int}})}{\cos(\alpha/2^n)}$. So we have

$\{S(\Omega_{n-1}^{ext})\} = \left\{\dfrac{S(\Omega_n^{\text{int}})}{\cos(\alpha/2^n)}\right\} = \dfrac{\{S(\Omega_n^{\text{int}})\}}{\{\cos(\alpha/2^n)\}}$. The sequence $\{S(\Omega_n^{\text{int}})\}$ goes to A, and the sequence $\{\cos(\alpha/2^n)\}$ goes to 1.

(**Comment**: α is a fixed angle, then the sequence $\{\alpha/2^n\} \to 0^o$ and the function $\cos\beta$ is continuous at 0^o, then $\{\cos(\alpha/2^n)\} \to \cos(0^o) = 1$).

Then, according to the basic properties of limits,

$$\left\{\frac{S(\Omega_n^{\text{int}})}{\cos(\alpha/2^n)}\right\}\to\frac{A}{1}\Rightarrow\{S(\Omega_{n-1}^{ext})\}\to A\Leftrightarrow\{S(\Omega_n^{ext})\}\to A.$$

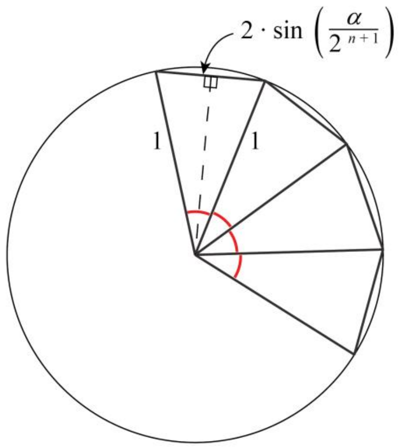
So, both internal and external sequences go to the same limit A, these sequences are the sequences of areas of internal and external figures (for the initial α-sector).

In the “Area construction” (Book 1) we have proved the assertion5 **[2-nd criterion of measurability]** from which follows that the area of α-sector is also equal to $A\equiv S(\alpha)$.

Let’s sum up: for any angle $\alpha\in[0^o,360^o]$, an α-sector is a measurable figure and it’s area $S(\alpha)$ can be determined as a limit of the sequence $\{2^{n-1}\cdot\sin(\alpha/2^n)\}=\{S(\Omega_n^{\text{int}})\}$, or as a limit of the sequence $\{2^n\cdot tg(\alpha/2^{n+1})\}=\{S(\Omega_n^{ext})\}$.

Let’s consider now any internal figure Ω_n^{int}, it defines n equal chords on the circle **[pict4]**. These chords are equal because they are bases of equal isosceles triangles.
In order to calculate such chord we can take any isosceles triangle with angle $\dfrac{\alpha}{2^n}$ and draw it’s height, it divides the triangle into two equal right triangles with angles $\dfrac{\alpha}{2\cdot 2^n}=\dfrac{\alpha}{2^{n+1}}$, so the base is

$$\sin\left(\frac{\alpha}{2^{n+1}}\right)+\sin\left(\frac{\alpha}{2^{n+1}}\right)=2\cdot\sin\left(\frac{\alpha}{2^{n+1}}\right).$$



pict.4

And there are 2^n such chords in total, then the sum of all chords of the internal figure Ω_n^{int} is $\Sigma_n=2^n\cdot 2\cdot\sin\left(\dfrac{\alpha}{2^{n+1}}\right)=2^{n+1}\cdot\sin\left(\dfrac{\alpha}{2^{n+1}}\right)$ **[V]**.

And the sequence of internal figures $\{\Omega_n^{\text{int}}\}$ defines the sequence $\{\Sigma_n\}$ of sum of chords.

Assertion6. The sequence $\{\Sigma_n\}$ of sum of the chords converges, and it’s limit is 2 times greater than the area $S(\alpha)$.

Proof. We have $\{\Sigma_n\}=\{2^{n+1}\cdot\sin(\alpha/2^{n+1})\}$ and $\{S(\Omega_n^{\text{int}})\}=\{2^{n-1}\cdot\sin(\alpha/2^n)\}$, we know that $\{S(\Omega_n^{\text{int}})\}\to S(\alpha)$. Let’s take the term of the sequence $\{S(\Omega_n^{\text{int}})\}$ with number $n+1$.

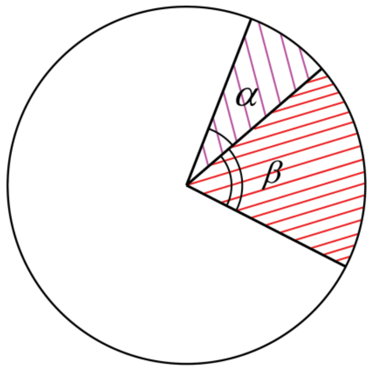
So $S(\Omega_{n+1}^{\text{int}})=2^n\cdot\sin(\alpha/2^{n+1})$ and we have $\Sigma_n=2^{n+1}\cdot\sin(\alpha/2^{n+1})$, then $\Sigma_n=2\cdot S(\Omega_{n+1}^{\text{int}})$, and we have $\{\Sigma_n\}=2\cdot\{S(\Omega_{n+1}^{\text{int}})\}$ and the sequence $\{S(\Omega_{n+1}^{\text{int}})\}$ goes to $S(\alpha)$, then $\{\Sigma_n\}$ goes to $2\cdot S(\alpha)$.

So, for any $\alpha\in[0^o,360^o]$ the numbers $L(\alpha),S(\alpha)$ are concrete real numbers.

Def. For any α-sector, the limit of the sequence $\{\Sigma_n\}$ is called a length of α-arc, and we denote it $L(\alpha)$. We got the formula $L(\alpha)=2\cdot S(\alpha)$ (remember that we operate on the unit circle Ω).

Assertion7 [Additivity]. For any angles $\alpha, \beta \in [0^o, 180^o]$ their sum $\alpha + \beta \in [0^o, 360^o]$ defines an $\alpha + \beta$-sector. And $S(\alpha) + S(\beta) = S(\alpha + \beta)$ and $L(\alpha) + L(\beta) = L(\alpha + \beta)$.

Proof. Let's fix any angles $\alpha, \beta \in [0^o, 180^o]$. We could use the formulas that we obtained above to prove this assertion, but there is a more simple way **[pict5]**. Let's build a central angle α and a central angle β right next to it. So we have an α-sector and a β-sector which form together an $\alpha + \beta$-sector. Both α, β-sectors do not have any common internal points, then, by additivity of area: $S(\alpha) + S(\beta) = S(\alpha + \beta)$ **[T]**. We showed above that for any angle $\gamma \in [0^o, 360^o]$ we have $L(\gamma) = 2S(\gamma)$, in particular it is true for the angles $\alpha, \beta, (\alpha + \beta)$. Let's multiply by 2 both sides of **[T]**, we will get $L(\alpha) + L(\beta) = L(\alpha + \beta)$.



pict.5

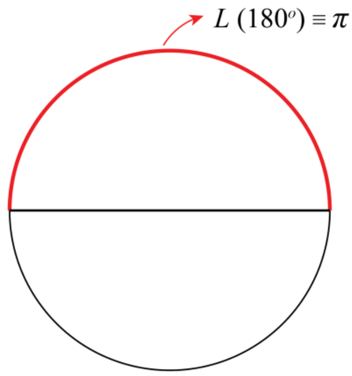
Consequence1. Let we have some α-sector (where $\alpha \in [0^o, 360^o]$).
Let's divide it into $\alpha_1, \alpha_2 \ldots \alpha_n$-sectors. Then the area of α-sector is a sum of areas of $\alpha_1, \alpha_2 \ldots \alpha_n$-sectors, and the length of α-arc is a sum of lengths of $\alpha_1, \alpha_2 \ldots \alpha_n$-arcs.

Def: Let's take any 180^o-sector, it defines a 180^o-arc **[pict6]**. The length of that arc $L(180^o)$ must be denoted by the letter π, so $L(180^o) \equiv \pi$.

A length of any arc is a limit of a concrete sequence, in this case we have $\pi = \lim_{n \to \infty} 2^{n+1} \cdot \sin\left(\frac{180^o}{2^{n+1}}\right) = L(180^o)$ **[V]**.

From here follows that the number π can be calculated as a limit of the sequence $\left\{2^{n+1} \cdot \sin\left(\frac{180^o}{2^{n+1}}\right)\right\}$.

In order to get a very good approximation of π we can calculate some concrete term $2^{n+1} \cdot \sin\left(\frac{180^o}{2^{n+1}}\right)$ for some big value of n.



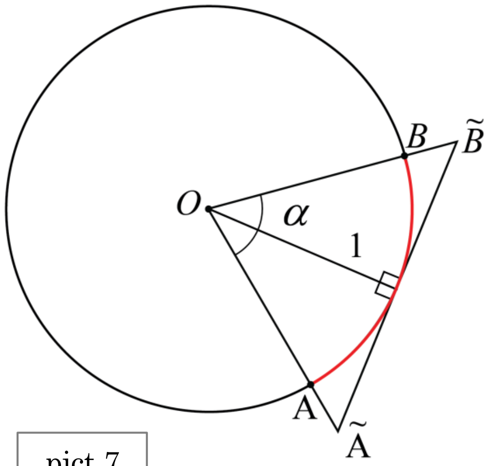
pict.6

Practical calculations show that $\pi \approx 3{,}1415$.

Consequence2. The full circle can be represented as two 180^o-arcs, then (consequence1) the length of the circle is 2π. The area of a half-circle is two times less than it's length, so it is $\pi / 2$ (because $L(\alpha) = 2S(\alpha)\ \forall \alpha$). The area of the full circle is π. By using simple ideas (as above): the length of 30^o arc is $\pi/6$, the length of 45^o arc is $\pi/4$, the length of 60^o arc is $\pi/3$.

Consequence3. For any $\alpha < \beta \in [0^o, 360^o]$ we have $L(\alpha) < L(\beta)$ and $S(\alpha) < S(\beta)$.
It immediately follows from the assertion7. Let's fix any $\alpha < \beta$, then the angle $(\beta - \alpha)$ is a positive angle. Let's build an α-sector and a $(\beta - \alpha)$-sector right next to it, together they form a β-sector and, by additivity, (assertion7) we have $S(\beta) = S(\alpha) + S(\beta - \alpha) \Rightarrow S(\beta) > S(\alpha)$. Now we can multiply by 2 both sides, then we get $L(\alpha) < L(\beta)$.

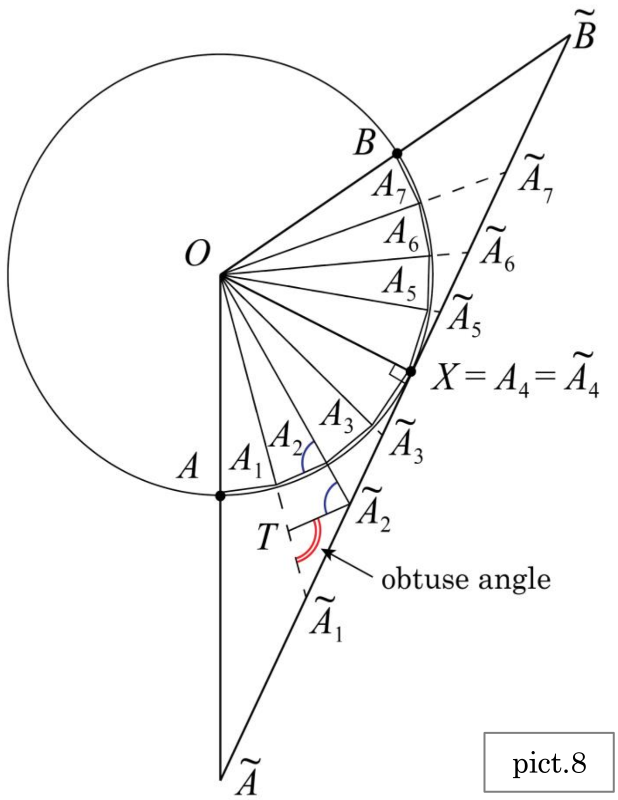
Auxiliary3. For any angle $\alpha \in (0^o, 180^o)$ the length of α-arc is not greater than the length of the segment $\tilde{A}\tilde{B} = 2tg(\alpha/2)$ **[pict7]**. Where $\tilde{A}, \tilde{B}$ are the points of intersection of a tangent line (which passes through the central point of α-arc) and rays OA, OB, where $\angle AOB = \alpha$.



pict.7

Proof. Let's fix any central angle $\alpha \in (0^o, 180^o)$.
Let OX is an angle bisector and l_X is a tangent line at the point X. Any internal figure Ω_n^{int} is symmetrical with respect to OX (because in order to build such figure, the angle α must be divided into 2^n equal parts, so there will be a symmetry with respect to OX).

Let's consider all chords of the internal figure Ω_n^{int}. The figure Ω_n^{int} defines the points $A, A_1, A_2, A_3 \ldots. A_k, B$ on the circle (in fact the number $k = 2^n - 1$ but we don't need it). Let's build the rays **[pict8]** $OA, OA_1, OA_2, OA_3 \ldots. OA_k, OB$, these rays intersect the tangent line l_X (which passes through X) at some points $\tilde{A}, \tilde{A}_1, \tilde{A}_2, \tilde{A}_3 \ldots. \tilde{A}_k, \tilde{B}$.
We will show that each chord is strictly less than the appropriate segment on l_X.
So $AA_1 < \tilde{A}\tilde{A}_1,\ A_1A_2 < \tilde{A}_1\tilde{A}_2,\ A_2A_3 < \tilde{A}_2\tilde{A}_3$ ...**[G]**.



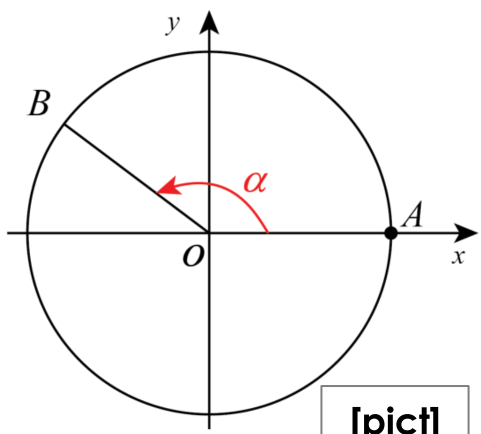
pict.8

Let's consider for example the chord A_1A_2, this chord is a base of the isosceles triangle ΔOA_1A_2, then $\angle OA_2A_1$ is an acute angle. Let's build the line $\tilde{A}_2T \parallel A_1A_2$, the angle $\angle O\tilde{A}_2T$ is equal to $\angle OA_2A_1$, so it is also an acute angle. Let's notice that $\angle O\tilde{A}_2\tilde{A}_1$ is an obtuse angle, because it is an exterior angle of the right triangle $\Delta OX\tilde{A}_2$ (where $\angle OX\tilde{A}_2 = 90^o$). So $\angle O\tilde{A}_2T$ is acute and $\angle O\tilde{A}_2\tilde{A}_1$ is obtuse, it means that $\tilde{A}_2T$ lies inside

the triangle $\Delta O\tilde{A}_2\tilde{A}_1$. As $\Delta O\tilde{A}_2 T$ is an isosceles triangle, the angle $\angle OT\tilde{A}_2$ is acute, then the adjacent angle $\angle\tilde{A}_1 T\tilde{A}_2$ is obtuse, then $\tilde{A}_1\tilde{A}_2$ is the greatest side in $\Delta\tilde{A}_1 T\tilde{A}_2$ (because it lies in front of an obtuse angle), in particular $\tilde{A}_1\tilde{A}_2 > \tilde{A}_2 T$ and obviously $\tilde{A}_2 T \ge A_1A_2$ (because $\tilde{A}_2 T$ and A_1A_2 are bases of similar isosceles triangles with the similarity coefficient $O\tilde{A}_2 / OA_2 \ge 1$), then $\tilde{A}_1\tilde{A}_2 > A_1A_2$.

So we have **[G]** $AA_1 < \tilde{A}\tilde{A}_1$, $A_1A_2 < \tilde{A}_1\tilde{A}_2$, $A_2A_3 < \tilde{A}_2\tilde{A}_3$ .... The sum of all chords AA_1, A_1A_2, A_2A_3 .... is Σ_n (the limit of such sums is the length of α-arc). And the sum of all segments AA_1, A_1A_2, A_2A_3 .... is $\tilde{A}\tilde{B}$. From **[G]** we have $\Sigma_n < \tilde{A}\tilde{B}$, so every term of the sequence $\{\Sigma_n\}$ is less than $\tilde{A}\tilde{B}$, then the limit of this sequence is not greater than $\tilde{A}\tilde{B}$, it means that $L(\alpha)$ is not greater than $\tilde{A}\tilde{B}$, so $L(\alpha) \le \tilde{A}\tilde{B}$. And finally, the segment $\tilde{A}\tilde{B}$ **[pict7]** consists of 2 equal cathetuses of right triangles, with the common hypotenuse 1 and angles $\alpha/2$, then $\tilde{A}\tilde{B} = 2tg(\alpha/2)$. Everything is proved.

Until this moment, when we were talking about an α-sector, we meant virtually any α-sector of a unit circle. Let's agree that from now on any central angle α must be built in the standard way, as we do it in trigonometry (Book I , "Angles"). So, any angle $\alpha \in [0^o, 360^o]$ must be drawn as an angle $\angle BOA$, where B is moved along the circle in the counterclockwise direction from A **[pict]**. And from now on, any α-sector is a concrete figure, which we get after we build the central angle $\alpha \in [0^o, 360^o]$ according to the standard rules. For any α-sector, the numbers $L(\alpha), S(\alpha)$ are uniquely defined. And we can say now that $L(\alpha), S(\alpha)$ are concrete functions on the segment $[0^o, 360^o]$.

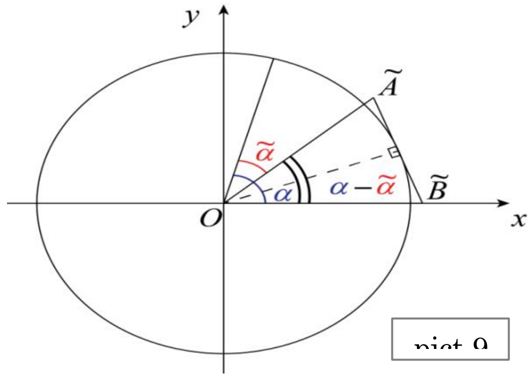


[pict]

From the consequence3 we know that for any $\alpha < \beta \in [0^o, 360^o]$ we have $S(\alpha) < S(\beta)$ and $L(\alpha) < L(\beta)$. So, the functions $L(\alpha), S(\alpha)$ are both strictly increasing on $[0^o, 360^o]$.

Assertion8. The function $L(\alpha)$ is continuous on $[0^o, 180^o)$.

Proof. We will use the estimation from the auxiliary3 here. Let's show that $L(\alpha)$ is continuous at any point $\tilde{\alpha} \in [0^o, 180^o)$. We fix an arbitrary $\tilde{\alpha} \in [0^o, 180^o)$. **If** $\alpha > \tilde{\alpha}$, then $L(\alpha) - L(\tilde{\alpha}) = L(\alpha - \tilde{\alpha})$ (the equality $L(\alpha) - L(\tilde{\alpha}) = L(\alpha - \tilde{\alpha}) \,||\, \alpha > \tilde{\alpha}$ immediately follows from the **additivity** of L). Obviously $(\alpha - \tilde{\alpha}) \in (0^o, 180^o)$, then (auxiliary3) $L(\alpha - \tilde{\alpha}) \le \tilde{A}\tilde{B}$, where $\tilde{A}\tilde{B}$



pict 9

is a segment of a tangent line which passes through the middle point of $(\alpha-\widetilde{\alpha})$-arc **[pict9]**.

And $\widetilde{A}\widetilde{B}=2tg\dfrac{(\alpha-\widetilde{\alpha})}{2}$. So we have $L(\alpha-\widetilde{\alpha})\le 2\cdot tg\dfrac{(\alpha-\widetilde{\alpha})}{2}$,

we can rewrite it $|L(\alpha)-L(\widetilde{\alpha})|\le 2\cdot tg\dfrac{(\alpha-\widetilde{\alpha})}{2}\Rightarrow$

$\Rightarrow 0<|L(\alpha)-L(\widetilde{\alpha})|\le 2\cdot tg\dfrac{(\alpha-\widetilde{\alpha})}{2}$ **[1]** for $\alpha>\widetilde{\alpha}$.

If we assume $\widetilde{\alpha}>\alpha$, we will get the similar inequality $0<|L(\widetilde{\alpha})-L(\alpha)|\le 2\cdot tg\dfrac{(\widetilde{\alpha}-\alpha)}{2}$ **[2]**.

We can unite the estimations **[1]** and **[2]** and write $0<|L(\alpha)-L(\widetilde{\alpha})|\le 2\cdot tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}$ **[E].**

Here we will use the squeeze theorem for functions. Obviously $\lim_{\alpha\to\widetilde{\alpha}} 0=0$ and if we show that also $\lim_{\alpha\to\widetilde{\alpha}} 2\cdot tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}=0$, then there must be $\lim_{\alpha\to\widetilde{\alpha}}|L(\alpha)-L(\widetilde{\alpha})|=0$ and it means that $|L(\alpha)-L(\widetilde{\alpha})|$ is infinitely small when $\alpha\to\widetilde{\alpha}$, then $L(\alpha)-L(\widetilde{\alpha})$ is infinitely small when $\alpha\to\widetilde{\alpha}$, then $\lim_{\alpha\to\alpha} L(\alpha)=L(\widetilde{\alpha})$ and $L(\alpha)$ is continuous at $\alpha=\widetilde{\alpha}$. Everything is proved.

Let's show that $2\cdot tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}$ **[F]** is infinitely small when $\alpha\to\widetilde{\alpha}$. Let's consider at first $tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}$ as a composite function $\mu(\varphi(\alpha))$, where $\varphi(\alpha)\equiv\dfrac{|\alpha-\widetilde{\alpha}|}{2}$ and $\mu(\beta)\equiv tg\beta$.

Obviously $\varphi(\alpha)$ is continuous at $\widetilde{\alpha}$ and $\mu(\beta)\equiv tg\beta$ is continuous at $\varphi(\alpha)=0^o$, then the composite function $\mu(\varphi(\alpha))$ is continuous at $\widetilde{\alpha}$, it means that

$\lim_{\alpha\to\widetilde{\alpha}}\mu(\varphi(\alpha))=\mu(\varphi(\widetilde{\alpha}))\Leftrightarrow\lim_{\alpha\to\widetilde{\alpha}} tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}=tg\dfrac{|\widetilde{\alpha}-\widetilde{\alpha}|}{2}=0$, then $tg\dfrac{|\alpha-\widetilde{\alpha}|}{2}$ is infinitely small when $\alpha\to\widetilde{\alpha}$, then **[F]** is infinitely small when $\alpha\to\widetilde{\alpha}$.

Consequence4. $L(\alpha)$ is continuous on $[0^o,360^o]$.

We already know that $L(\alpha)$ is continuous on $[0^o,180^o)$. Let's consider $L(\alpha)$ on $[180^o,270^o]$.

For any $\widetilde{\alpha}\in[180^o,270^o]$ there exist the unique $\alpha\in[0^o,90^o]$ such that $\widetilde{\alpha}=\alpha+180^o$ and according to additivity: $L(\widetilde{\alpha})=L(\alpha)+L(180^o)\Leftrightarrow L(\widetilde{\alpha})=L(\alpha)+\pi$. So the function $L(\alpha)$ on $[180^o,270^o]$ differs from $L(\alpha)$ on $[0^o,90^o]$ only by a constant π. Then $L(\alpha)$ is continuous on $[180^o,270^o]$.

In the exactly similar way we can show that $L(\alpha)$ is continuous on $[270^o,360^o]$.

After we sum up our results and conclude that $L(\alpha)$ is continuous on $[0^o,360^o]$.

Similarly, as $L(\alpha)$ is strictly increasing on $[0^o,180^o)$, then it is strictly increasing on $[0^o,360^o]$.

Then $L(\alpha)$ is strictly increasing and continuous on $[0^o,360^o]$.

According to the **inverse function theorem**, $L:[0^o,360^o]\to[0,2\pi]$ is one-to-one mapping, and there exist the inverse mapping $\alpha(L):[0,2\pi]\to[0^o,360^o]$ which is also continuous and strictly increasing.

The function $L(\alpha)$ takes the angle $\tilde{\alpha}\in[0^o,360^o]$ and compares to it the length $\tilde{L}\equiv L(\tilde{\alpha})$ of $\tilde{\alpha}$ -arc. The function $\alpha(L)$ makes the inverse action, it takes the number $\tilde{L}\in[0,2\pi]$ and compares the angle $\tilde{\alpha}\equiv\alpha(\tilde{L})\in[0^o,360^o]$ to it such that the length of $\tilde{\alpha}$ -arc is $\tilde{L}$.

Let's extend our function L on $[0^o,+\infty^o)$. For any angle $\alpha\ge 360^o$ there exist the unique representation $\alpha=360^o\cdot k+\tilde{\alpha}\,\|\,k\in\mathrm{N},\ \tilde{\alpha}\in[0^o,360^o)$ and we define: $L(\alpha)\equiv 2\pi\cdot k+L(\tilde{\alpha})$.

Geometrical meaning [1]. By definition, any angle $\alpha\ge 360^o$ we represent like $\alpha=360^o\cdot k+\tilde{\alpha}$ and we must understand such angle (by definition) as a combination of k full rotations around the circle, starting from A in the counterclockwise direction and some angle $\tilde{\alpha}<360^o$. According to our definition, for any $\alpha\ge 360^o$ the length of α -arc is the total length of k full circles (i.e., $2\pi\cdot k$) plus the length of $\tilde{\alpha}$ -arc, where $\tilde{\alpha}<360^o$.

Now $L(\alpha)$ is defined on $[0^o,+\infty)$, this set can be divided into segments $[360^o\cdot k,\ 360^o\cdot(k+1)]\ \|\,k\in\mathrm{Z},\ k\ge 0$. On any concrete segment $[360^o\cdot k,\ 360^o\cdot(k+1)]$ the function $L(\alpha)$ differs from the function $L(\alpha)$ on $[0^o,\ 360^o]$ only by a constant $2\pi k$.

Then $L(\alpha)$ is also strictly increasing and continuous on every segment $[360^o\cdot k,\ 360^o\cdot(k+1)]$.

Every segment $[0^o,\ 360^o\cdot n]$ can be represented as a union of segments $[0^o,\ 360^o],\ [360^o\cdot 1,\ 360^o\cdot 2]\ \ldots.\ [360^o\cdot(n-1),\ 360^o\cdot n]$, on every of which $L(\alpha)$ is strictly increasing and continuous, then $L(\alpha)$ is strictly increasing and continuous on $[0^o,\ 360^o\cdot n]$.

Theorem1. $L(\alpha)$ is strictly increasing and continuous on $[0^o,+\infty^o)$ and $L(\alpha)$ is one-to-one mapping $[0^o,+\infty^o)\to[0,+\infty)$. The inverse mapping $\alpha(L):[0,+\infty)\to[0^o,+\infty^o)$ is also strictly increasing and continuous.

Proof. This theorem is very simple, but anyway, let's give the detail proof of it.

[Increase of $L(\alpha)$]. Let's fix any $\alpha_1<\alpha_2\in[0^o,+\infty^o)$ there exist the segment $[0^o,\ 360^o\cdot n]$ that contains both these points α_1,α_2. As $L(\alpha)$ is strictly increasing on $[0^o,\ 360^o\cdot n]$, then $L(\alpha_1)<L(\alpha_2)$, then $L(\alpha)$ is strictly increasing on $[0^o,+\infty^o)$.

[Continuity of $L(\alpha)$]. We just need to show that $L(\alpha)$ is continuous at any point $\tilde{\alpha} \in [0^o, +\infty^o)$, so let's fix now any $\tilde{\alpha} \in [0^o, +\infty^o)$, by definition, $L(\alpha)$ is continuous at $\tilde{\alpha}$ if for any sequence $\{\alpha_n\} \to \tilde{\alpha} \,||\, \{\alpha_n\} \in [0^o, +\infty^o)$ we have $\{L(\alpha_n)\} \to L(\tilde{\alpha})$.

Let's fix an arbitrary sequence $\{\tilde{\alpha}_n\} \to \tilde{\alpha} \,||\, \{\tilde{\alpha}_n\} \in [0^o, +\infty^o)$, as this sequence converges, it must be bounded, then there exist the segment $[0^o,\ 360^o \cdot n]$ which contains all the terms of the sequence $\{\tilde{\alpha}_n\}$ and $\tilde{\alpha}$. As $L(\alpha)$ is continuous on $[0^o,\ 360^o \cdot n]$, then it is continuous at $\tilde{\alpha} \in [0^o,\ 360^o \cdot n]$, and for the sequence $\{\tilde{\alpha}_n\} \to \tilde{\alpha} \,||\, \{\tilde{\alpha}_n\} \in [0^o,\ 360^o \cdot n]$ we have $\{L(\tilde{\alpha}_n)\} \to L(\tilde{\alpha})$.

Then $L(\alpha)$ is continuous on $[0^o, +\infty^o)$.

Let's show that $L(\alpha)$ is one-to-one mapping $[0^o, +\infty^o) \to [0, +\infty)$.

[Step1] we need to show that for any number $b \in [0, +\infty)$ there exist some $\alpha_b \in [0^o, +\infty^o)$ such that $L(\alpha_b) = b$.

Let's consider the numbers $2\pi, b$. By the **Archimedes axiom** there exist $\tilde{n} \in \mathrm{N}$ such that $b < 2\pi \cdot \tilde{n}$. Let's consider the segment $[0^o,\ 360^o \cdot \tilde{n}]$, as we noticed above, $L(\alpha)$ is strictly increasing and continuous on $[0^o,\ 360^o \cdot \tilde{n}]$. According to the **inverse function theorem**, $L(\alpha)$ makes one-to-one mapping $[0^o,\ 360^o \cdot \tilde{n}] \to [0,\ 2\pi \cdot \tilde{n}]$, as $b \in (0,\ 2\pi \cdot \tilde{n})$, then there exist some $\alpha_b \in [0^o,\ 360^o \cdot \tilde{n}]$ such that $L(\alpha_b) = b$.

[Step2] Let's show that $L(\alpha)$ doesn't "glue together" elements of $[0^o, +\infty^o)$. Really as L is strictly increasing on $[0^o, +\infty^o)$, then for any $\alpha_1 \neq \alpha_2 \in [0^o, +\infty^o)$ we have $L(\alpha_1) \neq L(\alpha_2)$.

From the **[Step1]** and **[Step2]** follows that $L(\alpha)$ is one-to-one mapping $[0^o, +\infty^o) \to [0, +\infty)$. Then the inverse mapping $\alpha(L) : [0, +\infty) \to [0^o, +\infty^o)$ is defined.

Let's show that $\alpha(L) : [0, +\infty) \to [0^o, +\infty^o)$ is also strictly increasing and continuous.

As $L(\alpha)$ is strictly increasing and continuous on every segment $[0^o,\ 360^o \cdot \tilde{n}]$, then the inverse function $\alpha(L)$ must be strictly increasing and continuous on every image $[0,\ 2\pi \cdot \tilde{n}]$ of such segment. So, for any $\tilde{n} \in \mathrm{N}$ the function $\alpha(L)$ is strictly increasing and continuous on $[0,\ 2\pi \cdot \tilde{n}]$, it's very easy to deduce from here that $\alpha(L)$ is strictly increasing and continuous on $[0, +\infty)$

(the proof consists of exactly the same steps as **[Increase of $L(\alpha)$]**, **[Continuity of $L(\alpha)$]**, but now these steps must be done for $\alpha(L)$). Everything is proved.

Let's finally extend our function L on $(-\infty^o,+\infty^o)$.

Now $L(\alpha)$ is defined for any $\alpha\in[0^o,+\infty^o)$, let's define $L(\alpha)$ for any $\alpha\in(-\infty^o,0^o]$.

For any negative angle $\alpha\in(-\infty^o,0^o]$. We have $\alpha=-|\alpha|$, and we define: $L(\alpha)\equiv //\,by\ def\,// \equiv -L(|\alpha|)$.

So, for any negative angle α, the length of α-arc is defined like the number which is opposite to the length of $|\alpha|$-arc.

Geometrical meaning [2]. Let's build any negative angle $-\alpha\in[-360^o,0^o]$ on the unit circle. Then the value $L(\alpha)$ is a minus length of the arc which is opposite to the negative angle $-\alpha$.

For any $\alpha\le-360^o$ there exist the unique representation $\alpha=-360^o\cdot k-\tilde{\alpha}\ \|\ k\in\mathrm{N},\ -\tilde{\alpha}\in(-360^o,0^o]$, and this angle (by definition) must be understanded as k full rotations around the circle in the clockwise direction and some negative angle $-\tilde{\alpha}$. And according to our definition, $L(\alpha)$ is a minus total length of k full circles, minus the length of the arc which is opposite to the negative angle $-\tilde{\alpha}$.

Now $L(\alpha)$ is defined for any angle $\alpha\in(-\infty^o,+\infty^o)$.

As $L(\alpha)$ is strictly increasing and continuous on $[0^o,+\infty^o)$, then $L(\alpha)$ is strictly **decreasing** and **continuous** on $(-\infty^o,0^o]$. As $L(\alpha)$ on $[0^o,+\infty^o)$ is one-to-one mapping $[0^o,+\infty^o)\to[0,+\infty)$, then $L(\alpha)$ on $(-\infty^o,0^o]$ is one-to-one mapping $(-\infty^o,0^o]\to(-\infty,0]$.

As $L(\alpha)$ is continuous on $(-\infty^o,0^o]$ and on $[0^o,+\infty^o)$, then $L(\alpha)$ is continuous everywhere $(-\infty^o,+\infty^o)$.

Let's also extend the function $\alpha(L)$. Now it is defined on $[0,+\infty)$. We extend it in the exactly similar way as above, for any $L\in(-\infty,0]$ we have $L=-|L|$, and we define $\alpha(L)\equiv //\,by\ def\,// \equiv -\alpha(|L|)$.

As $\alpha(L)$ is strictly increasing and continuous on $[0,+\infty)$, then $\alpha(L)$ is **strictly decreasing** and **continuous** on $(-\infty,0]$. As $\alpha(L)$ on $[0,+\infty)$ is one-to-one mapping $[0,+\infty)\to[0^o,+\infty^o)$, then $\alpha(L)$ on $(-\infty,0]$ is one-to-one mapping $(-\infty,0]\to(-\infty^o,0^o]$.

As $\alpha(L)$ is continuous on $(-\infty,0]$ and on $[0,+\infty)$, then $\alpha(L)$ is continuous everywhere $(-\infty,+\infty)$

Exercise3. Show that $\alpha(L)\,\|\,L\in(-\infty,+\infty)$ is an inverse function for $L(\alpha)\,\|\,\alpha\in(-\infty^o,+\infty^o)$.

(use that $\alpha(L)$ on $[0,+\infty)$ is already an inverse function for $L(\alpha)$ on $[0^o,+\infty^o)$ and use also the simple rules, according to which we extended these functions).

Let's sum up, now $L(\alpha)$ is defined on $(-\infty^o,+\infty^o)$ and the inverse function $\alpha(L)$ is defined on $(-\infty,+\infty)$. Let's notice that when L goes from $-\infty$ to 0, the angle value $\alpha(L)$ goes from $-\infty^o$ to 0^o, when L goes from 0 to $+\infty$, the angle value $\alpha(L)$ goes from 0^o to $+\infty^o$. So in general, when L goes from $-\infty$ to $+\infty$, the angle value $\alpha(L)$ goes from $-\infty^o$ to $+\infty^o$.

Let's consider trigonometric functions $\sin\alpha, \cos\alpha \,||\, \alpha \in (-\infty^o,+\infty^o)$ **[R1]**. Each of these functions must be turned into a composite function $\sin(\alpha(L)),\ \cos(\alpha(L)) \,||\, L \in (-\infty,+\infty)$ **[R2]**.

When L goes from $-\infty$ to $+\infty$, the angle value $\alpha(L)$ goes from $-\infty^o$ to $+\infty^o$.
And the new representation does not "restrict" the angle value $\alpha(L)$, and now the variable L (which is a real number) is our main variable.

Def: For any angle $\alpha \in (-\infty^o,+\infty^o)$, the number $L(\alpha)$ (which can be understanded as a length of α-arc ,**Geometrical meaning [1],[2]**) is called a **radian measure** of α.

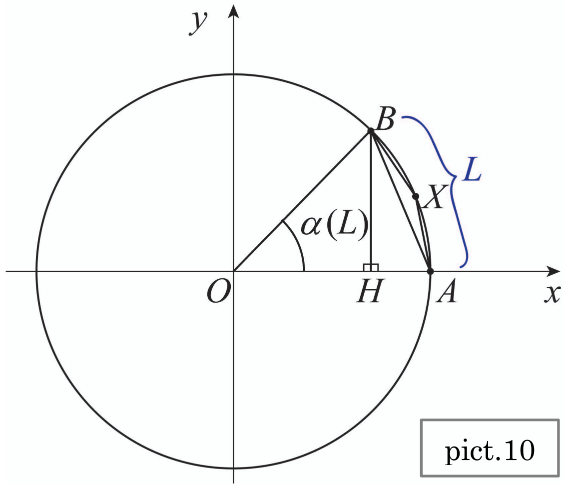
And we can say "$L(\alpha)$ is α in radians". For example, "π is 180^o in radians" and "2π is 360^o in radians" and "$\pi/3$ is 60^o in radians" and "$\pi/2$ is 90^o in radians".
In general, any number $L \in (-\infty,+\infty)$ can be called "L radians". And now our functions $\sin(\alpha(L)),\ \cos(\alpha(L)) \,||\, L \in (-\infty,+\infty)$ are the functions of a radian argument L.

Assertion9. $\sin(\alpha(L)),\ \cos(\alpha(L))$ are both continuous on $(-\infty,+\infty)$.

Proof. The function $\alpha(L)$ is continuous on $(-\infty,+\infty)$, the values of $\alpha(L)$ on $(-\infty,+\infty)$ belong to $(-\infty^o,+\infty^o)$. And the function $\sin\alpha$ is continuous on $(-\infty^o,+\infty^o)$, then (continuity of a composite function) the function $\sin(\alpha(L))$ is continuous on $(-\infty,+\infty)$. And similarly for $\cos(\alpha(L))$.

Exercise4. What are the domains of composite functions $tg(\alpha(L)),\ ctg(\alpha(L))$?
Explain why each of these functions is continuous on it's domain.

Theorem2. $\lim_{L\to 0}\dfrac{\sin\alpha(L)}{L}=1$



pict.10

Proof. Let's consider the right limit $\lim_{L\to 0+}\dfrac{\sin\alpha(L)}{L}$.

We fix any positive number $L\in(0,\pi/2)$, it defines the unique angle $\alpha(L)\in(0^o,90^o)$ such that the length of $\alpha(L)$-arc is L **[pict10]**. We have shown earlier the estimation $L\le 2tg\dfrac{\alpha(L)}{2}$ **[ES]** (auxiliary3).

We need to find the lower estimate of L. Let's denote $\alpha(L)\equiv\angle AOB$, we draw the perpendicular BH, obviously $BH<AB$ and $AB<AX+XB$, where X is a central point of $\alpha(L)$-arc, let's notice that the sum $AX+XB$ is the first term Σ_1 of the sequence $\{\Sigma_n\}$, the limit of this sequence is the length L of $\alpha(L)$-arc. The sequence $\{\Sigma_n\}$ is always strictly increasing (we can understand it by simple geometrical reasoning, or just by using the connection between $\{\Sigma_n\}$ and a strictly increasing sequence $S(\Omega_n^{\text{int}})$). Then we have: $\sin\alpha(L)=BH<AB<AX+XB=\Sigma_1<L$.

Let's take **[ES]**, then we have $\sin\alpha(L)<L\le 2tg\dfrac{\alpha(L)}{2}$, let's divide all the sides by $\sin\alpha(L)>0$,

then $1<\dfrac{L}{\sin\alpha(L)}\le\dfrac{2tg\frac{\alpha(L)}{2}}{\sin\alpha(L)}=\dfrac{2\cdot\sin\frac{\alpha(L)}{2}\Big/\cos\frac{\alpha(L)}{2}}{2\sin\frac{\alpha(L)}{2}\cos\frac{\alpha(L)}{2}}=\dfrac{1}{\left(\cos\frac{\alpha(L)}{2}\right)^2}$ and we have

$1<\dfrac{L}{\sin\alpha(L)}\le\dfrac{1}{\left(\cos\frac{\alpha(L)}{2}\right)^2}$ from here follows that $\left(\cos\dfrac{\alpha(L)}{2}\right)^2\le\dfrac{\sin\alpha(L)}{L}<1$ **[Z].**

And we will use the squeeze theorem for functions here. Obviously $\lim_{L\to 0+}1=1$.

Let's show that $\lim_{L\to 0+}\left(\cos\dfrac{\alpha(L)}{2}\right)^2=1$, if it's true, then from **[Z]** there must be

$\lim_{L\to 0+}\dfrac{\sin\alpha(L)}{L}=1$. We will show that the ordinary limit $\lim_{L\to 0}\left(\cos\dfrac{\alpha(L)}{2}\right)^2=1$ (then the right limit $\lim_{L\to 0+}\left(\cos\dfrac{\alpha(L)}{2}\right)^2$ exists and equals 1). Let's consider $\cos\dfrac{\alpha(L)}{2}$ as a composite function $\mu(\varphi(L))$, where $\varphi(L)\equiv\dfrac{\alpha(L)}{2}$ and $\mu(\beta)\equiv\cos\beta$. The function $\varphi(L)\equiv\dfrac{\alpha(L)}{2}$ is continuous at the point $\tilde{L}=0$ (because $\alpha(L)$ is continuous everywhere) and the function $\mu(\beta)=\cos\beta$

is continuous at $\varphi(0)=0^o$, then the composite function $\mu(\varphi(L))$ is continuous at $\tilde{L}=0$, and we have $\lim_{L\to 0}\mu(\varphi(L))=\mu(\varphi(0)) \Leftrightarrow \lim_{L\to 0}\cos\frac{\alpha(L)}{2}=\cos\frac{\alpha(0)}{2}=\cos 0^o=1$.

We showed that $\lim_{L\to 0}\left(\cos\frac{\alpha(L)}{2}\right)=1$, then, obviously $\lim_{L\to 0}\left(\cos\frac{\alpha(L)}{2}\right)^2=1^2=1$.

Therefore $\lim_{L\to 0+}\frac{\sin\alpha(L)}{L}=1$. The function $\frac{\sin\alpha(L)}{L}\equiv\varphi(L)$ is an even function which is defined on $L\in(-\infty,0)\cup(0,+\infty)$, and it's graph has a symmetry with respect to Oy.

It follows from $\varphi(-L)=\varphi(L)$ for any $L\neq 0$. For any such function: if one of the limits $\lim_{L\to 0+}\varphi(L)$, $\lim_{L\to 0-}\varphi(L)$ exist, then the other limit also exists and these limits are equal.

So there must be $\lim_{L\to 0-}\frac{\sin\alpha(L)}{L}=1$. The function $\frac{\sin\alpha(L)}{L}$ has both left and right limits at the point $\tilde{L}=0$, and these limits are equal to 1, then the ordinary limit $\lim_{L\to 0}\frac{\sin\alpha(L)}{L}$ exists, and it is also equal to 1. Everything is proved.

In mathematics all the functions $\sin,\cos,tg,ctg$ are usually considered as functions of a radian argument L. And instead of writings $\sin(\alpha(L))$, $\cos(\alpha(L))$, $tg(\alpha(L))$, $ctg(\alpha(L))$ people write just $\sin x$, $\cos x$, tgx, $ctgx$, where x is assumed to be a radian variable L. So we assume $x\equiv L$ and we just neglect the letter α. Let's notice that all the basic trigonometric formulas are still true. Any formula $\sin 2\alpha=2\sin\alpha\cos\alpha$, which is true for any angle α, can be rewritten for $\alpha\equiv\alpha(L)$, so $\sin 2\alpha(L)=2\sin\alpha(L)\cdot\cos\alpha(L)$ and we replace L by x and neglect the letter α, so we have $\sin 2x=2\sin x\cdot\cos x$. And the same can be done with any other trigonometric formula.

The theorem2 can be written as $\lim_{x\to 0}\frac{\sin x}{x}=1$, it is a very important limit in mathematics.
We will use it to obtain the derivatives of all trigonometric functions. From now on we consider all trigonometric functions as functions of a radian argument $x\in(-\infty,+\infty)$, such approach is extremely important, only now we can define inverse trigonometric functions, later, by using radians, we will be able to provide Taylor serieses for trigonometric functions and etc.

Inverse trigonometric functions

Let's consider $\sin x \,\|\, x\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$, $\cos x \,\|\, x\in[0,\pi]$, $tgx \,\|\, x\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$, $ctgx \,\|\, x\in(0,\pi)$,

each of these functions is continuous and strictly increasing/decreasing on it's segment/interval. Then (according to the inverse function theorem) each of these functions is one-to-one mapping:

$$\sin x:\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\to[-1,1] \ \ || \ \ \cos x:[0,\pi]\to[-1,1] \ \ ||$$

$$tgx:\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\to(-\infty,+\infty) \ || \ \ ctgx:(0,\pi)\to(-\infty,+\infty)$$

and for each of these functions the inverse function is defined, which is also continuous and strictly increasing/decreasing on it’s domain.

The inverse functions are denoted like:

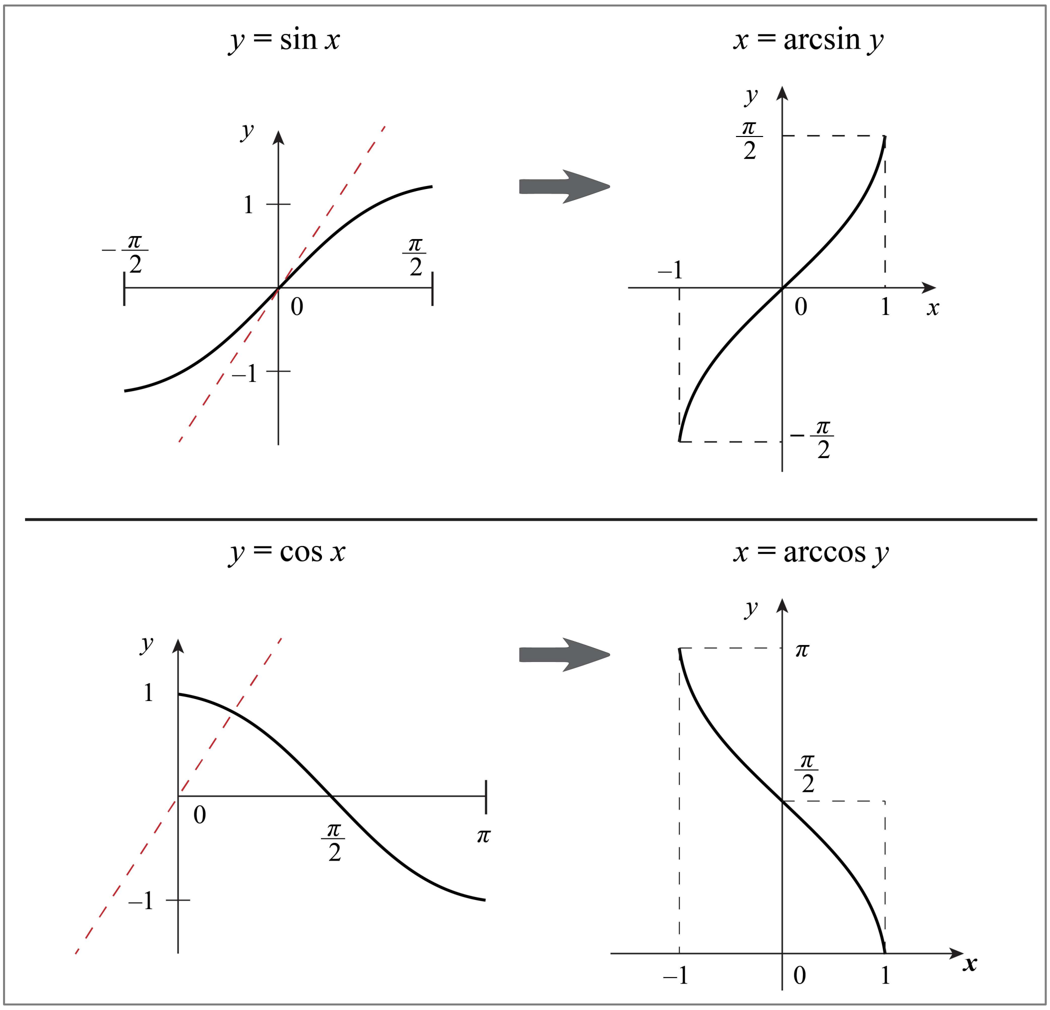
$$\arcsin y:[-1,1]\to\left[-\frac{\pi}{2},\frac{\pi}{2}\right] \ \ || \ \ \arccos y:[-1,1]\to[0,\pi] \ \ ||$$

$$arctgy:(-\infty,+\infty)\to\left(-\frac{\pi}{2},\frac{\pi}{2}\right) || \ \ arcctg\ y:(-\infty,+\infty)\to(0,\pi)$$

.

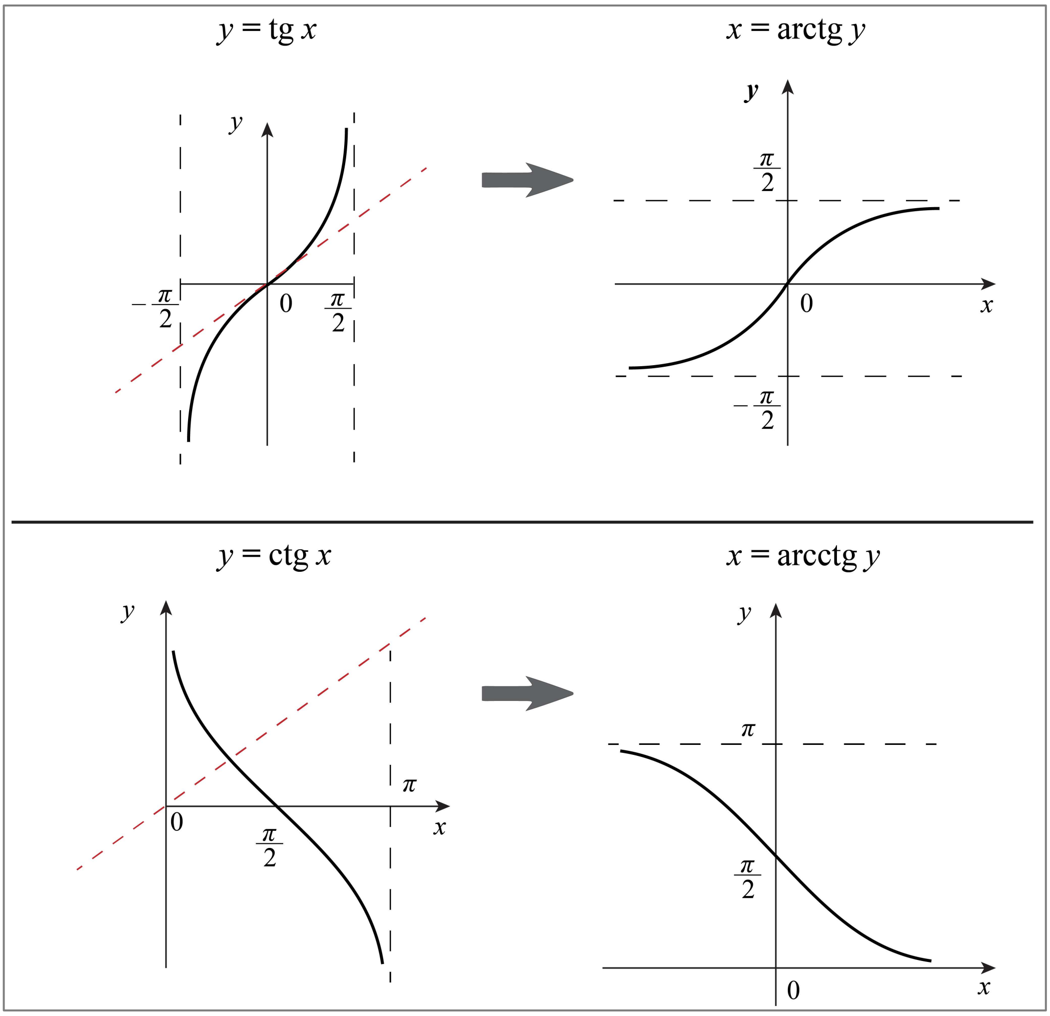
These functions have very simple meanings. For example, $\arcsin y$ is the angle $x\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ (in radians) such that $\sin x = y$. And $\mathrm{arctg}\ y$ is the angle $x\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ (in radians) such that $tgx = y$ and etc.

The graphs of the inverse trigonometric functions can be easily obtained from the graphs of the initial trigonometric functions, we just need to reflect each of these graphs over the line $y = x$ (red line) **[pict1], [pict2]**.



pict1

pict2



Exercise1. If we pay attention to the pictures, it's easy to notice that for any $x \in \left(0, \frac{\pi}{2}\right)$ the line $y = x$ lies above the graph $y = \sin x$, but below the graph $y = \mathrm{tg} x$, it means that for any $x \in \left(0, \frac{\pi}{2}\right)$ there must be: $\sin x < x < \mathrm{tg} x$. Show that this equality is really true (it's easy to do by using the estimations that we got during the constructions of radians).

For Reader's practice:

[1] Show that for any $y \in [-1,1]$ we have $\arcsin y + \arccos y = \dfrac{\pi}{2}$.

[2] Prove that $\sin(2\arcsin x) = 2x\sqrt{1-x^2}$.

[3] Given that $\arccos y + \arccos z + \arccos h = \pi$, prove that $y^2 + z^2 + h^2 + 2yzh = 1$.

[4] Solve the equation $(\mathrm{arctg} x)^2 + (\arccos x)^2 = \dfrac{5\pi^2}{8}$. **Answer:** $x = -1$.

[5] Prove that $\arcsin\dfrac{12}{13} + \arccos\dfrac{4}{5} + arctg\dfrac{63}{16} = \pi$.

Power function

In any multiplicative group G, an integer power $a^m \| m \in Z$ of any element $a \in G$ is defined.

In particular, for any real number $a \in R$ the real number $a^m \in R$ is defined (m is any integer number). We want to define the meaning of symbols like $(\sqrt{2})^{\sqrt{2}}$, $3^{\sqrt{5}}, \sqrt{18}^{1/5}$ and etc., i.e., we want to throw away the restriction for the power m "to be integer".

We have already defined what is $\sqrt[k]{a} \| a \in R, \ \forall k \in \mathrm{N}$. And we showed that for any $a \geq 0$ and for any $k \in \mathrm{N}$ the real number $\sqrt[k]{a}$ is defined, now we define: $a^{1/k} \equiv \sqrt[k]{a}$.

If we want to define $a^m \| a \in R, \ m \in R$, we have to make the restriction for a to be positive (at least non negative). Really, even now we see that for any negative number $a<0$ the number $a^{1/4} = \sqrt[4]{a}$ is not defined, by definition $\sqrt[4]{a}$ is such real number b that $b^4 = a$, but in any case $b^4 \geq 0$, and we have $a<0$, then $b^4 \neq a$, and the element $a^{1/4}$ is not defined.

So, from now on, a is always a positive real number $a>0$.

According to the standard properties of integer power, in any commutative multiplicative group we have: **[A]** $a^m \cdot a^n = a^{m+n}$, **[B]** $(a \cdot b)^m = a^m \cdot b^m$, **[C]** $a^{mn} = (a^m)^n = (a^n)^m$, **[D]** $(a^m)^{-1} = (a^{-1})^m$.

In particular it is true for any $a, b \in R, \ m, n \in Z$ (in the last case **[D]** there must be $a \neq 0$).

Auxiliary1. If for some **positive** real numbers a, b we have $a^d = b^d$ for some natural d, then $a=b$.

Comment. The field of real numbers R is an ordered field, and it's very easy to get Auxiliary1 if we use the Property4 **[C]** of ordered rings/fields, Book1, page43.

Def1. Let $a \in R, \ a>0$ is any real number and $q \in Q$ is any rational number. There exist the representation $q = \dfrac{m}{k} \| m \in Z, \ k \in \mathrm{N}$. We define $a^q = a^{m/k} \equiv // by\ def // \equiv (a^{1/k})^m \| m \in Z$.

Assertion1. The **def1** is correct. The main point here is that any rational number can be represented as a quotient of an integer and a natural in many ways, let $q = \dfrac{m}{k} = \dfrac{\tilde{m}}{\tilde{k}} \| m, \tilde{m} \in Z, \ k, \tilde{k} \in \mathrm{N}$ we need to show that $a^{m/k}$ is equal to $a^{\tilde{m}/\tilde{k}}$.

The main idea: $a^{m/k}$ and $a^{\tilde{m}/\tilde{k}}$ are positive numbers, we will show that they are equal if they are raised to the natural power $k \cdot \tilde{k}$, then, according to auxiliary1, these numbers are equal.

Let's calculate $(a^{m/k})^{k \cdot \tilde{k}}$, the number $a^{m/k} = // by\ def // = (a^{1/k})^m$, then $(a^{m/k})^{k \cdot \tilde{k}} = ((a^{1/k})^m)^{k \cdot \tilde{k}}$, here the element $a^{1/k}$ is some real number and $m, k, \tilde{k}$ are integers, so we can rewrite $((a^{1/k})^k)^{m \cdot \tilde{k}} = a^{m \cdot \tilde{k}}$ **[1]** and in the exactly similar way we can get $(a^{\tilde{m}/\tilde{k}})^{k \cdot \tilde{k}} = a^{\tilde{m} \cdot k}$ **[2]**.

It's easy to see that the expressions **[1]** and **[2]** are equal, really $\dfrac{m}{k}=\dfrac{\tilde{m}}{\tilde{k}} \Rightarrow m\tilde{k}=\tilde{m}k$, then $a^{m\cdot\tilde{k}}=a^{\tilde{m}\cdot k}$ and **[1]** and **[2]** are equal. Then (**the main idea**) we have $a^{m/k}=a^{\tilde{m}/\tilde{k}}$.

Assertion2. For any positive real numbers $a,b>0$ and for any $q,q_1,q_2\in Q$ we have:

[A] $(a\cdot b)^q=a^q\cdot b^q$, **[B]** $a^{q_1+q_2}=a^{q_1}\cdot a^{q_2}$, **[C]** $a^{q_1\cdot q_2}=(a^{q_1})^{q_2}=(a^{q_2})^{q_1}$.

Let's prove **[A]**. We fix any rational number $q\in Q$ and it's representation $q=\dfrac{m}{k}\,\|\, m\in \mathrm{Z},\ k\in \mathrm{N}$.

Then $(a\cdot b)^q=(a\cdot b)^{m/k}=/\,by\ def\,/=((a\cdot b)^{1/k})^m$ and $a^q\cdot b^q=a^{m/k}\cdot b^{m/k}=(a^{1/k})^m\cdot(b^{1/k})^m$.
We already have the property **[A]** for any integer power m, so $(a^{1/k})^m\cdot(b^{1/k})^m=(a^{1/k}\cdot b^{1/k})^m$.
And we need to show that numbers $(a^{1/k}\cdot b^{1/k})^m$ and $((a\cdot b)^{1/k})^m$ are equal. It's obviously enough to prove that $a^{1/k}\cdot b^{1/k}$ and $(a\cdot b)^{1/k}$ are equal. Let's raise these numbers to the natural power k, if we get equal results, then (auxiliary1) these numbers are equal.
So $(a^{1/k}\cdot b^{1/k})^k=(a^{1/k})^k\cdot(b^{1/k})^k=a\cdot b$ and $((a\cdot b)^{1/k})^k=a\cdot b$.

Consequence from **[A]**. For any positive $a,b>0$ and any $q\in Q$ we have $\left(\dfrac{a}{b}\right)^q=\dfrac{a^q}{b^q}$.

Proof. This equality is equivalent to $\left(\dfrac{a}{b}\right)^q\cdot b^q=a^q$ the left part, according to **[A]**, can be simplified as $\left(\dfrac{a}{b}\right)^q\cdot b^q=\left(\dfrac{a}{b}\cdot b\right)^q=a^q$.

For **[B]**,**[C]** we fix an arbitrary pair of rational numbers $q_1,q_2\in Q$ and their representations $q_1=\dfrac{m_1}{k_1},\ q_2=\dfrac{m_2}{k_2}\,\|\, m_1,m_2\in \mathrm{Z},\ k_1,k_2\in \mathrm{N}$.

Let's prove **[B]**. $a^{q_1+q_2}=a^{m_1k_2+k_1m_2/k_1k_2}=(a^{1/k_1k_2})^{m_1k_2+k_1m_2}$ and

$a^{q_1}\cdot a^{q_2}=a^{m_1/k_1}\cdot a^{m_2/k_2}=(a^{1/k_1})^{m_1}\cdot(a^{1/k_2})^{m_2}$. We need to show that $(a^{1/k_1k_2})^{m_1k_2+k_1m_2}$ **[3]** is equal to $(a^{1/k_1})^{m_1}\cdot(a^{1/k_2})^{m_2}$ **[4]**. We have here some positive real numbers which are raised to some integer powers. Let's show that both these numbers are equal if they are raised to the natural power $k_1\cdot k_2$. So $\left((a^{1/k_1k_2})^{m_1k_2+k_1m_2}\right)^{k_1\cdot k_2}=\left((a^{1/k_1k_2})^{k_1\cdot k_2}\right)^{m_1k_2+k_1m_2}=a^{m_1k_2+k_1m_2}$ and
$\left((a^{1/k_1})^{m_1}\cdot(a^{1/k_2})^{m_2}\right)^{k_1\cdot k_2}=(a^{1/k_1})^{k_1\cdot k_2\cdot m_1}\cdot(a^{1/k_2})^{k_2\cdot k_1\cdot m_2}=\left((a^{1/k_1})^{k_1}\right)^{k_2\cdot m_1}\cdot\left((a^{1/k_2})^{k_2}\right)^{k_1\cdot m_2}=$
$=a^{k_2\cdot m_1}\cdot a^{k_2\cdot m_1}=a^{m_1k_2+k_1m_2}$. We see that numbers **[3]** and **[4]** are equal if they are raised to the natural power k_1k_2, then these numbers are equal.

Let's prove **[C]**. We have $q_1 \cdot q_2 = \frac{m_1}{k_1} \cdot \frac{m_2}{k_2} = \frac{m_1 \cdot m_2}{k_1 \cdot k_2}$, then

$a^{q_1 \cdot q_2} = a^{m_1 \cdot m_2 / k_1 \cdot k_2} = / by\ def / = (a^{1/k_1 k_2})^{m_1 \cdot m_2} = ((a^{1/k_1 k_2})^{m_1})^{m_2}$ **[5]** and

$(a^{q_1})^{q_2} = (a^{m_1/k_1})^{m_2/k_2} = (((a^{1/k_1})^{m_1})^{1/k_2})^{m_2}$ **[6]**. If we show that $((a^{1/k_1})^{m_1})^{1/k_2}$ **[7]** and

$(a^{1/k_1 k_2})^{m_1}$ **[8]** are equal, then expressions **[5]** and **[6]** are equal, and it is exactly what we need. Both **[7]** and **[8]** are positive real numbers, if they become equal after they raised to the natural power k_2, then these numbers are equal. So **[7]** $(((a^{1/k_1})^{m_1})^{1/k_2})^{k_2} = (a^{1/k_1})^{m_1} = a^{m_1/k_1}$ and

[8] $((a^{1/k_1 k_2})^{m_1})^{k_2} = (a^{1/k_1 k_2})^{m_1 k_2} = a^{m_1 k_2 / k_1 k_2}$. And obviously a^{m_1/k_1} and $a^{m_1 k_2 / k_1 k_2}$ are equal, because the fractions $\frac{m_1}{k_1}$ and $\frac{m_1 \cdot k_2}{k_1 \cdot k_2}$ are equal (in the assertion1 we have proved that if $\frac{m}{k} = \frac{\tilde{m}}{\tilde{k}}$, then $a^{m/k} = a^{\tilde{m}/\tilde{k}}$).

Theorem1. If a **positive** sequence of real numbers $\{x_n\}$ goes to a **positive** a, then for any rational number $q \in Q$ we nave $\{(x_n)^q\} \to a^q$.

Proof. Let's fix any representation $q = \frac{m}{k} \| m \in \mathrm{Z},\ \ k \in \mathrm{N}$.

[Step1] $\{x_n\} \to a$, then $\{(x_n)^{1/k}\} \to a^{1/k}$. In "radians" (assertion1) we showed that for any $k \in \mathrm{N}$ the function $f(x) \equiv \sqrt[k]{x}$ continuous on $[0,+\infty)$. The sequence $\{x_n\}$ and a both belong to $[0,+\infty)$, as $\{x_n\} \to a$, then $\{f(x_n)\} \to f(a) \Leftrightarrow \{(x_n)^{1/k}\} \to a^{1/k}$.

[Step2] Let's take the representation $q = \frac{m}{k}$. If m is a natural number, then according to the simpliest properties of sequence limits, we have

$\{(x_n)^{1/k}\} \to a^{1/k} \Rightarrow \{((x_n)^{1/k})^m\} \to (a^{1/k})^m \Leftrightarrow \{(x_n)^{m/k}\} \to a^{m/k} \Leftrightarrow \{(x_n)^q\} \to a^q$.

Let's return to representation $q = \frac{m}{k}$. If $m = 0$, then $q = 0$ and $\{(x_n)^q\} = \{(x_n)^0\} = 1,1,1....$ and $a^q = a^0 = 1$, then $\{(x_n)^q\} \to a^q$.

Let's return to representation $q = \frac{m}{k}$. If m is a negative number, then $m = -|m|$, where $|m| \in \mathrm{N}$.

The number $\frac{|m|}{k}$ is a positive rational number and we already know that $\{(x_n)^{|m|/k}\} \to a^{|m|/k}$ - here we have a positive sequence that goes to some positive number, then:

$$\left\{\frac{1}{(x_n)^{|m|/k}}\right\} \to \frac{1}{a^{|m|/k}} \Leftrightarrow \{(x_n)^{-|m|/k}\} \to a^{-|m|/k} \Leftrightarrow \{(x_n)^{m/k}\} \to a^{m/k} \Leftrightarrow \{(x_n)^q\} \to a^q.$$

Assertion3. For any $a \in R \,||\, a > 0$ we have $\lim_{n\to\infty} \ a^{1/n} = 1$.

Proof. Let $a > 1$. Let's denote $\forall n \,||\ \ \beta_n = \sqrt[n]{a} - 1 > 0 \Rightarrow \sqrt[n]{a} = 1 + \beta_n$. Then, according to the Bernoulli inequality, (Book1,page 118) we have $\Rightarrow \beta_n \le \dfrac{a-1}{n}$, then finally $0 < \beta_n \le \dfrac{a-1}{n}$ which is equivalent to $0 < \sqrt[n]{a} - 1 \le \dfrac{a-1}{n}$, let's use here the squeeze theorem for sequences, we will get:
$\lim_{n\to\infty} (\sqrt[n]{a} - 1) = 0 \Rightarrow \lim_{n\to\infty} \sqrt[n]{a} = 1$.

If $a = 1$, then $\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \sqrt[n]{1} = 1$. And the last case: $0 < a < 1 \Rightarrow a = \dfrac{1}{b} \,||\, b > 1$, we already know that $\lim_{n\to\infty} \sqrt[n]{b} = 1$. From here $\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \sqrt[n]{\dfrac{1}{b}} = \dfrac{\lim_{n\to\infty} \sqrt[n]{1}}{\lim_{n\to\infty} \sqrt[n]{b}} = \dfrac{1}{1} = 1$.

Consequence1. For any $a \in R \,||\, a > 0$ we have $\lim_{n\to\infty} \ a^{-1/n} = 1$.

Proof. $\lim_{n\to\infty} \ a^{-1/n} = \lim_{n\to\infty} \ \dfrac{1}{a^{1/n}} = \dfrac{\lim_{n\to\infty} 1}{\lim_{n\to\infty} a^{1/n}} = \dfrac{1}{1} = 1$.

Assertion4. For any $a > 1$ the function a^q is strictly increasing on Q, i.e.,
$\forall q_1 < q_2 \in Q \Rightarrow a^{q_1} < a^{q_2}$.

Proof. [part1] $a > 1$ then for any pair of consecutive integer numbers $n, (n+1)$ we have $a^n < a^{n+1}$.
Really: a^n is a positive number and $a^n \cdot 1 < [as\ 1 < a] < a^n \cdot a = a^{n+1}$.

Consequence2. For any integer numbers $n_1 < n_2$ we have $a^{n_1} < a^{n_2}$.

[part2] Let $q_1 = \dfrac{m_1}{k_1},\ q_2 = \dfrac{m_2}{k_2} \,||\, m_1, m_2 \in Z,\ k_1, k_2 \in \mathrm{N}$ and $\dfrac{m_1}{k_1} < \dfrac{m_2}{k_2}$. Both a^{m_1/k_1} and a^{m_2/k_2} are positive numbers, if they are both raised to the natural power $k_1 \cdot k_2$ and we see that one of these numbers is greater than the other, then it was true from the very beginning (**auxiliary1**).
As we showed above: $(a^{m_1/k_1})^{k_1 \cdot k_2} = a^{m_1 \cdot k_2}$ **[9]** and $(a^{m_1/k_1})^{k_1 \cdot k_2} = a^{m_1 \cdot k_2}$ **[10]**, we know that $\dfrac{m_1}{k_1} < \dfrac{m_2}{k_2} \Rightarrow m_1 k_2 < m_2 k_1$ (these are integer numbers), then from the consequence2 we have $a^{m_1 \cdot k_2} < a^{m_1 \cdot k_2}$. So **[9]** is less than **[10]**. Then $a^{m_1/k_1} < a^{m_2/k_2}$.

Assertion5. For any $a > 1$ we have $\lim_{q\to 0 || q\in Q} a^q = 1$.

Proof. Let's fix an arbitrary rational sequence $\{q_n\} \to 0$, if we show that $\{a^{q_n}\} \to 1$, then $\lim_{q\to 0 || q\in Q} a^q = 1$. We fix an arbitrary $\varepsilon > 0$. According to the assertion3 and consequence1 from

it: $\lim_{n\to\infty} a^{-1/n}=1$ and $\lim_{n\to\infty} a^{1/n}=1$. For $\bar{\varepsilon}>0$ there exist the natural $\bar{k}$ such that $\forall n>\bar{k}$ we have $|a^{-1/n}-1|<\bar{\varepsilon}$ and $|a^{1/n}-1|<\bar{\varepsilon}$, let's rewrite: $1-\bar{\varepsilon}<a^{-1/n}<1+\bar{\varepsilon}$ **[E1]** and **[E2]** $1-\bar{\varepsilon}<a^{1/n}<1+\bar{\varepsilon}$. As $a>1$, then for any $n\in\mathrm{N}$ we have $1<a^{1/n}$, from here follows that $\frac{1}{1}>\frac{1}{a^{1/n}}\Leftrightarrow 1>a^{-1/n}$, then $a^{-1/n}<1<a^{1/n}$ **[E3].** From **[E1]** and **[E2]** and **[E3]** we have: $1-\bar{\varepsilon}<a^{-1/n}<1<a^{1/n}<1+\bar{\varepsilon}\,\|\,\forall n>\bar{k}$. We need to fix here any natural number $\bar{n}$ such that $1-\bar{\varepsilon}<a^{-1/\bar{n}}<1<a^{1/\bar{n}}<1+\bar{\varepsilon}$ **[E4]**.

Let's take now the sequence $\{q_n\}$. We take $\delta=1/\bar{n}$, as $\{q_n\}\to 0$, then, starting from some number $\bar{m}$, there must be $|q_n-0|<1/\bar{n}\Rightarrow -1/\bar{n}<q_n<1/\bar{n}$ - from this inequality and assertion4 follows that $a^{-1/\bar{n}}<a^{q_n}<a^{1/\bar{n}}$, then from **[E4]** we have $1-\bar{\varepsilon}<a^{q_n}<1+\bar{\varepsilon}\Leftrightarrow |a^{q_n}-1|<\bar{\varepsilon}$ and it is true for any $n>\bar{m}$.

We had started from an arbitrary positive $\bar{\varepsilon}>0$, and we deduced that there exist some $\bar{m}$, such that $\forall n>\bar{m}$ we have $|a^{q_n}-1|<\bar{\varepsilon}$, it means that $\{a^{q_n}\}\to 1$. Everything is proved.

We are ready to define the power function in the case $a>1$.

Def2. $a>1$ is a constant and $x\in R$ is any real number. Let $\{q_n\}\to x$ is any sequence of rational numbers that goes to x, then the sequence $\{a^{q_n}\}$ converges and it's limit must be taken as a value $a^x\equiv /\,by\ def\,/\equiv \lim_{n\to\infty} a^{q_n}$.

Assertion6. The **def2** is correct. At first, any real number x can be represented as a limit of some rational sequence (Book1, page 94, **[Example L]**). Secondly, if some rational sequence $\{q_n\}\to x$, then $\{a^{q_n}\}$ converges, and for any sequence $\{q_n\}$ that goes to x, the sequence $\{a^{q_n}\}$ always goes to the same limit. Let's prove it. We fix any x and any sequence $\{q_n\}\to x$.

[1-st part] We will show that $\{a^{q_n}\}$ is fundamental.

The sequence $\{q_n\}$ converges, therefore it is bounded, then there exist some rational $\mathrm{M}>0:\ |q_n|<\mathrm{M}\ \forall n$. Let's fix an arbitrary $\bar{\varepsilon}>0$, according to the assertion5, $\lim_{q\to 0\|q\in Q} a^q=1$, then for the positive number $\bar{\varepsilon}/a^{\mathrm{M}}$ there exist $\bar{\delta}>0$ such that $\forall q:|\,q-0\,|<\bar{\delta}\Rightarrow |a^q-1|<\bar{\varepsilon}/a^{\mathrm{M}}$ **[H].**

And the final step: as $\{q_n\}$ converges, then $\{q_n\}$ is fundamental, then for $\bar{\delta}>0$ there exist k starting from which $\forall m,n>k$ we have $|q_m-q_n|<\delta$. Let's consider the absolute value $|a^{q_m}-a^{q_n}|$

for any $m,n>k$. We have: $\left|a^{q_m}-a^{q_n}\right|=\left|a^{q_n}\right|\cdot\left|a^{q_m-q_n}-1\right|$ **[S]**, here q_m-q_n is a rational number, and it's absolute value is less than δ, then from **[H]** we have $\left|a^{q_m-q_n}-1\right|<\bar{\varepsilon}/a^{\mathrm{M}}$.

And $\left|a^{q_n}\right|=a^{q_n}<[assertion\,4\;\; q_n<\mathrm{M}]<a^{\mathrm{M}}$. Then from **[S]** we have $\left|a^{q_m}-a^{q_n}\right|<\bar{\varepsilon}$.

Let's sum up: we had started from an arbitrary positive $\bar{\varepsilon}>0$ and we deduced that there exist k such that $\forall m,n>k$ we have $\left|a^{q_m}-a^{q_n}\right|<\bar{\varepsilon}$. It means that $\{a^{q_n}\}$ is fundamental.

R is a complete field, then $\{a^{q_n}\}$ converges in R to some number, which must be denoted as a^x.

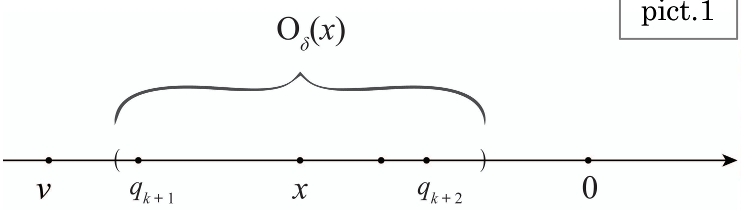
[2-nd part] Let's show that for any other rational sequence $\{v_n\}\to x$ the sequence $\{a^{v_n}\}$ goes to the same limit a^x (and therefore a^x is uniquely defined). Let's fix any other sequence $\{v_n\}\to x$, and we have $\{q_n\}\to x$. Let's consider the sequence $q_1,v_1,q_2,v_2,q_3,v_3...$, this sequence also goes to x. As we have shown in the [1-st part] the sequence $a^{q_1},a^{v_1},a^{q_2},a^{v_2},a^{q_3},a^{v_3}....$ converges to some limit. We already know that it's subsequence $a^{q_1},a^{q_2},a^{q_3},a^{q_4}$ .... goes to a^x. When a sequence converges to some limit, any it's subsequence converges to the same limit. Then $a^{q_1},a^{v_1},a^{q_2},a^{v_2},a^{q_3},a^{v_3}....$ may converge only to a^x (if we assume that it converges to some other number, we will immediately get a contradiction). Then it's subsequence a^{v_1},a^{v_2},a^{v_3} .... also goes to a^x.

So, for any $a>1$ and $x\in R$, the number a^x is defined. Therefore, for any concrete $a>1$ we have the function $f(x)\equiv a^x$. Let's explore the properties of this function.

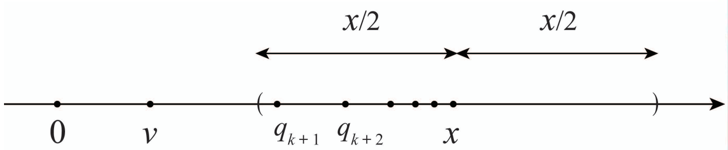
[1-st property]. For any $a>1$ the function a^x is positive everywhere $x\in(-\infty,+\infty)$. And for any positive x we have $a^x>1$.

Proof. Let's fix an arbitrary real number x and any rational sequence $\{q_n\}\to x$.
Let's fix some neighborhood $\mathrm{O}_\delta(x)$, it contains all the terms of $\{q_n\}$ starting from some number k. Let's fix any rational number v "on the left" of $\mathrm{O}_\delta(x)$ **[pict1]**, then we have $v<q_n \,||\, \forall n>k$.

pict.1



Then ([assertion4](#)) we have $a^v<a^{q_n}$, then obviously $a^v\le\lim_{n\to\infty}a^{q_n}=a^x$.
And for rational number v we obviously have $0<a^v$ (because a is positive), then there must be $0<a^x$. So a^x is positive everywhere $(-\infty,+\infty)$.



pict.2

Next, let's fix any positive x and any rational

sequence $\{q_n\} \to x$. Let's fix for example the neighborhood $O_\delta(x) \| \delta = x/2$ of x **[pict2]**.

$O_\delta(x)$ contains all the terms of the sequence $\{q_n\}$ starting form some number k.

Then $\exists k : \forall n > k$ we have $q_n \in O_\delta(x)$. Let's fix any rational v between 0 and $x/2$ **[pict2]**,

then $0 < v < q_n \| \forall n > k$.From here, according to the assertion4, we have

$a^0 < a^v < a^{q_n} \Leftrightarrow 1 < a^v < a^{q_n}$. From $a^v < a^{q_n} \| \forall n > k$ follows that $a^v \le \lim_{n\to\infty} a^{q_n} = a^x$.

Then $1 < a^v \le a^x \Rightarrow 1 < a^x$. Everything is proved.

[2-nd property]. For any $a, b > 1$ and for any real $x, x_1, x_2 \in R$ we have:

[A] $(a \cdot b)^x = a^x \cdot b^x$, **[B]** $a^{x_1} \cdot a^{x_2} = a^{x_1 + x_2}$ and **[C]** $(a^{x_1})^{x_2} = a^{x_1 \cdot x_2}$.

Proof. Let's fix any real x and any rational sequence $\{q_n\} \to x$. Then $a^x = \lim_{n\to\infty} a^{q_n}$ and $b^x = \lim_{n\to\infty} b^{q_n}$. The element $(a \cdot b)^x$ (by definition) is a limit of the sequence $(a \cdot b)^{q_n}$.

We know that **[A]** is true for any rational powers, so we can write $(a \cdot b)^{q_n} = (a)^{q_n} \cdot (b)^{q_n}$ and according to the standard properties of limits we have

$(a \cdot b)^x = \lim_{n\to\infty} (a \cdot b)^{q_n} = \lim_{n\to\infty} (a)^{q_n} \cdot (b)^{q_n} = \lim_{n\to\infty} (a)^{q_n} \cdot \lim_{n\to\infty} (b)^{q_n} = a^x \cdot b^x$.

Consequence from **[A]**. $\left(\dfrac{a}{b}\right)^x = \dfrac{a^x}{b^x} \| \begin{pmatrix} \forall a, b > 1 \\ \forall x \in R \end{pmatrix}$.

For the cases **[B]** and **[C]** we fix any rational sequences $\{q_n\} \to x_1$ and $\{v_n\} \to x_2$.

Let's prove **[B]**: $a^{x_1} = \lim_{n\to\infty} a^{q_n}$ and $a^{x_2} = \lim_{n\to\infty} a^{v_n}$, as $\{q_n\}$ and $\{v_n\}$ go to x_1 and x_2, then the rational sequence $\{q_n + v_n\}$ goes to $(x_1 + x_2)$ and $a^{x_1 + x_2} = \lim_{n\to\infty} a^{v_n + q_n}$. We know that **[B]** is true for any rational powers, and according to the standard properties of limits we have:

$\lim_{n\to\infty} a^{v_n + q_n} = \lim_{n\to\infty} (a^{v_n} \cdot a^{q_n}) = \lim_{n\to\infty} a^{v_n} \cdot \lim_{n\to\infty} a^{q_n} \Leftrightarrow a^{x_1 + x_2} = a^{x_1} \cdot a^{x_2}$.

Let's prove **[C]**: a^{x_1} is a limit of the sequence $\{a^{q_n}\}$, so $a^{x_1} = \lim_{n\to\infty} a^{q_n}$, then $(a^{x_1})^{x_2}$ is a limit of the sequence $(\lim_{n\to\infty} a^{q_n})^{v_1}$, $(\lim_{n\to\infty} a^{q_n})^{v_2}$, $(\lim_{n\to\infty} a^{q_n})^{v_3}$, $(\lim_{n\to\infty} a^{q_n})^{v_4}$ .... **[S1]**.

Next, the rational sequence $\{q_n \cdot v_n\}$ goes to $x_1 x_2$, then $a^{x_1 \cdot x_2}$ is a limit of the sequence $a^{q_1 \cdot v_1}, a^{q_2 \cdot v_2}, a^{q_3 \cdot v_3}, a^{q_4 \cdot v_4}$....... As **[C]** is true for any rational powers, then the last sequence can be rewritten as $(a^{q_1})^{v_1}, (a^{q_2})^{v_2}, (a^{q_3})^{v_3}, (a^{q_4})^{v_4}$... **[S2]**.

We need to show that sequences **[S1]** and **[S2]** have the same limit. Then we need to show that their difference is an infinitely small sequence.

Let's denote $\beta_k \equiv (a^{q_k})^{v_k} - (\lim_{n\to\infty} a^{q_n})^{v_k}$, then $\beta_k \equiv (a^{q_k})^{v_k} \cdot \left(1 - \dfrac{(\lim_{n\to\infty} a^{q_n})^{v_k}}{(a^{q_k})^{v_k}}\right)$.

According to the consequence from **[A]** (for rational powers) we have $\left(\dfrac{a}{b}\right)^q = \dfrac{a^q}{b^q} \| \begin{pmatrix} \forall a, b > 0 \\ \forall q \in Q \end{pmatrix}$.

Here we have $\lim_{n\to\infty} a^{q_n} > 0$ (follows from the [1-st property]) and also $a^{q_k} > 0$. Then we can rewrite: $\beta_k \equiv (a^{q_k})^{v_k} \cdot \left(1 - \left(\frac{\lim_{n\to\infty} a^{q_n}}{a^{q_k}}\right)^{v_k}\right)$ **[J]**. When $k \to \infty$ the sequence $\left\{\frac{\lim_{n\to\infty} a^{q_n}}{a^{q_k}}\right\}$ is positive and goes to 1, then (theorem1) the sequence $\left\{\left(\frac{\lim_{n\to\infty} a^{q_n}}{a^{q_k}}\right)^{v_k}\right\}$ goes to $1^{v_k} = 1$, then the sequence $\left\{1 - \left(\frac{\lim_{n\to\infty} a^{q_n}}{a^{q_k}}\right)^{v_k}\right\}$ goes to zero 0. And the sequence $\{(a^{q_k})^{v_k}\}$ goes to some real number (it actually goes to $a^{x_1 \cdot x_2}$, as we can notice above). Then the sequence **[J]** goes to zero 0, and the sequence $\{\beta_k\}$ (which is **[J]**) is infinitely small. Everything is proved.

[3-rd property]. For any $a > 1$ the function a^x is **strictly increasing** and **continuous** everywhere $x \in (-\infty, +\infty)$.

Proof. Let's fix an arbitrary pair of real numbers $x_1 < x_2$. Then $(x_2 - x_1) > 0$ is a positive real number. According to the [1-st property] we have $a^{x_2 - x_1} > 1$, let's multiply both part by $a^{x_1} > 0$ (it is positive according to the [1-st property]). Then
$a^{x_2 - x_1} \cdot a^{x_1} > 1 \cdot a^{x_1} \Leftrightarrow [2-nd\ property] \Leftrightarrow a^{(x_2 - x_1) + x_1} > a^{x_1} \Leftrightarrow a^{x_2} > a^{x_1}$, then a^x is **strictly increasing** on $(-\infty, +\infty)$.

Let's fix now an arbitrary point $x_0 \in (-\infty, +\infty)$. If we show that for any sequence $\{x_n\} \to x_0$ we have $\{a^{x_n}\} \to a^{x_0}$, then a^x is continuous at x_0 (by definition). Let's fix an arbitrary sequence $\{\bar{x}_n\} \to x_0$. For every term $\bar{x}_n$, we can find the rational number q_n such that $|\bar{x}_n - q_n| < 1/n$ **[L]**, from here immediately follows that $\{q_n\} \to x_0$. Then $a^{x_0} = \lim_{n\to\infty} a^{q_n}$ (by definition of a^{x_0}).
We want to show that the sequence $\{a^{\bar{x}_n}\}$ also goes to a^{x_0}. We already know that $\{a^{q_n}\}$ goes to a^{x_0}, then it's enough to show that the difference of these sequences $\{a^{\bar{x}_n}\} - \{a^{q_n}\}$ is an infinitely small sequence (then they go to the same limit). Let's consider $\{a^{\bar{x}_n} - a^{q_n}\}$. And let's estimate $a^{\bar{x}_n} - a^{q_n}$ **[P]**. From **[L]** we have $q_n - 1/n < \bar{x}_n < q_n + 1/n$, as a^x is strictly increasing, then $a^{q_n - 1/n} < a^{\bar{x}_n} < a^{q_n + 1/n}$, let's subtract a^{q_n} from all sides, in order to get the estimation of **[P]**:
$a^{q_n - 1/n} - a^{q_n} < a^{\bar{x}_n} - a^{q_n} < a^{q_n + 1/n} - a^{q_n}$ **[P1]** and now we will just use the squeeze theorem for sequences. Let's show that $\{a^{q_n + 1/n} - a^{q_n}\}$ is infinitely small, we represent $a^{q_n + 1/n} - a^{q_n} = a^{q_n}(a^{1/n} - 1)$.
When $n \to \infty$ the sequence $\{a^{1/n}\}$ goes to 1 (assertion3), then $\{a^{1/n} - 1\}$ goes to 0.
And the sequence $\{a^{q_n}\}$ goes to some real number (which is a^{x_0}), then the sequence $\{a^{q_n}(a^{1/n} - 1)\}$ goes to zero. So, the sequence $\{a^{q_n + 1/n} - a^{q_n}\}$ is infinitely small. Similarly, in **[P1]** the sequence

$\{a^{q_n - 1/n} - a^{q_n}\}$ is infinitely small. Then from **[P1]**, the sequence $\{a^{\bar{x}_n} - a^{q_n}\}$ is infinitely small. Everything is proved.

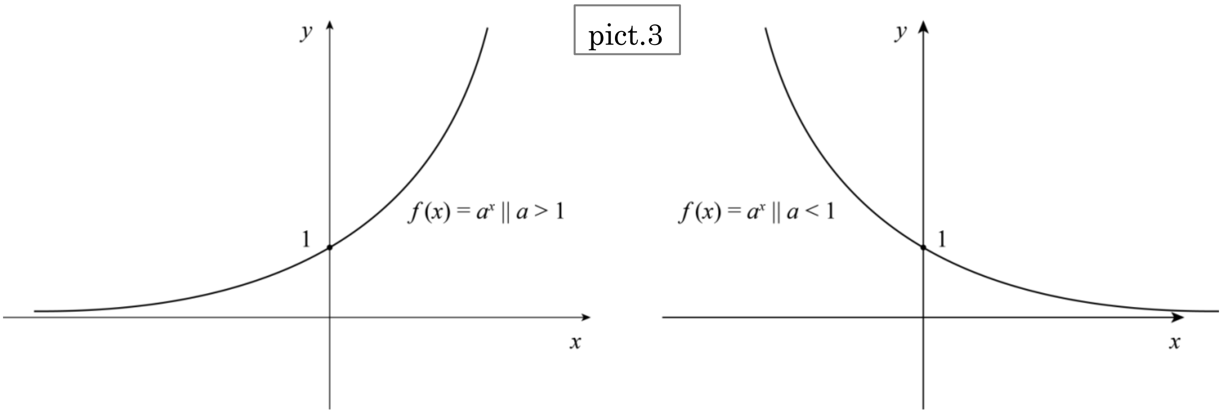
Let's extend the definition of a power function.

Def. For $a = 1$ we define $1^x \equiv 1 \| \forall x \in R$. For any $a \in (0,1)$ we take the representation $a = \dfrac{1}{b} \| b > 1$ (which is unique) and we define $a^x \equiv \dfrac{1}{b^x} \| \forall x \in R$.

This definition is correct, in the [1-st property] we showed that $b^x \| b > 1$ is positive for any $x \in (-\infty, +\infty)$, then $a^x \| a < 1$ is really defined on $(-\infty, +\infty)$.

From this definition immediately follows that for any concrete $a \in (0,1)$ the function a^x is strictly decreasing, continuous and strictly positive everywhere, and the [2-nd property] is true for this function. For $a = 1$ the power function is a constant function (and in this case it is also continuous, strictly positive everywhere and the [2-nd property] is true). Now the power function $a^x > 1$ is defined for every concrete $a > 0$ **[pict3]**. And we know it's properties.

pict.3



Logarithmic function

Let's fix any $a > 1$. We know that a^x is strictly increasing and continuous on $(-\infty, +\infty)$.

Exercise1. Show that $\lim_{x \to +\infty} a^x = +\infty$ and $\lim_{x \to -\infty} a^x = 0$.

Exercise2. Show that a^x is one-to-one mapping $(-\infty, +\infty) \to (0, +\infty)$.

Exercise3. By using the inverse function theorem show that there exist the inverse function $f^{-1}(x) : (0, +\infty) \to (-\infty, +\infty)$ which is also strictly increasing and continuous on $(0, +\infty)$.

Def. The inverse function from the exercise3 is called a logarithmic function and we denote it $\log_a x \equiv f^{-1}(x)$.

So, $\log_a x$ is defined on $(0, +\infty)$ and it is strictly increasing and continuous on $(0, +\infty)$.

The simple equality: $a^{\log_a x} = x \| \forall x > 0$ (which is exactly $f(f^{-1}(x)) = x$) shows us a very simple and important meaning of a logarithmic function: $\log_a x$ is the power, to which a must be raised in order to become equal to x.

When we see the writing $\log_a x$, we usually say that a is a base of the logarithm $\log_a x$. And $\log_a x$ can be called a "logarithm of x to the base a".

It's easy to prove that for any $a > 1, b > 1$ and for any $x_1, x_2 \in (0, +\infty)$ the next properties are true:
[1] $\log_a a^x = x$ **[2]** $\log_a(x_1 \cdot x_2) = \log_a(x_1) \cdot \log_a(x_2)$ **[3]** $\log_a x^b = b \cdot \log_a x \,||\, \forall b \in R$

[4] $\log_{a^b} x = (1/b) \cdot \log_a x \,||\, \forall b \in R,\ b \neq 0$ **[5]** $\log_a b \cdot \log_b a = 1$ **[6]** $\log_a x = \dfrac{\log_b x}{\log_b a}$.

We can also consider the power function a^x where $a \in (0,1)$ and we can perform exactly the same reasoning as above (the only difference here is that a^x is strictly decreasing, then $\log_a x$ must be also strictly decreasing), and we can get exactly the same properties **[1]-[5]** as above.

Def. The logarithm $\log_e x$, where the base is an exponent e, is called a "natural logarithm" and we denote it $\ln x \equiv \log_e x$.

Upper and lower limit

We have already defined the notion "subsequence", and we have proved that any bounded sequence has a convergent subsequence **[T]**. We have also noticed the next obvious thing: if a sequence converges to some limit, then any it's subsequence converges to the same limit.

Def. Let we have some sequence $\{x_n\}$ and it has a subsequence $\{x_{n_k}\}$ that goes to a. Then a is called a partial limit of the sequence $\{x_n\}$.

From now on we consider only bounded sequences $\{x_n\}$. For any bounded sequence, the set P, which consists of all partial limits of this sequence, is not empty. (We have proved **[T]** that P contains at least one number). And P is a bounded set. Really, $\{x_n\}$ is bounded, therefore all it's terms belong to some segment $[a,b]$ and any partial limit of this sequence must also belong to $[a,b]$. Then $\mathrm{P} \subset [a,b]$ and therefore the numbers $\inf \mathrm{P}$ and $\sup \mathrm{P}$ are defined.

Assertion1. The numbers $\inf \mathrm{P}$ and $\sup \mathrm{P}$ are also partial limits of $\{x_n\}$, i.e., there exist some subsequence $\{x_{n_k}\} \to \inf \mathrm{P}$ and some subsequence $\{x_{p_k}\} \to \sup \mathrm{P}$.

Proof. Let's fix any positive sequence $\{\varepsilon_k\} \to 0$. As $\inf \mathrm{P}$ is an infinum of P, the segment $[\inf \mathrm{P}, \inf \mathrm{P} + \varepsilon_1/2)$ contains some partial limit p_1, as p_1 is a limit of some subsequence, the set $(p_1 - \varepsilon_1/2, p_1 + \varepsilon_1/2)$ contains all the terms of that subsequence, starting from some number, let's just fix any term which belongs to $(p_1 - \varepsilon_1/2, p_1 + \varepsilon_1/2)$, it becomes the first term x_{n_1} of our subsequence. Obviously $|x_{n_1} - \inf \mathrm{P}| < \varepsilon_1$. The same thing must be done for the number ε_2, and we will find x_{n_2} such that $|x_{n_2} - \inf \mathrm{P}| < \varepsilon_2$ (It's important to notice here that we must choose $n_2 > n_1$)

and etc. As a result we have the subsequence $\{x_{n_k}\}: \left|x_{n_k} - \inf \mathrm{P}\right| < \varepsilon_k$ and $\{\varepsilon_k\} \to 0$, then $\{x_{n_k}\} \to \inf \mathrm{P}$. Similarly we can find the subsequence $\{x_{p_k}\} \to \sup \mathrm{P}$.

Def. The numbers $\inf \mathrm{P}$ and $\sup \mathrm{P}$ are called a lower and an upper limit of the sequence $\{x_n\}$ and we denote them as $\underline{\lim}_{n\to\infty} x_n \equiv \inf \mathrm{P}$ and $\overline{\lim}_{n\to\infty} x_n \equiv \inf \mathrm{P}$. From the assertion1 follows that we can say "$\underline{\lim}_{n\to\infty} x_n$ is a minimal among all partial limits of $\{x_n\}$" and "$\overline{\lim}_{n\to\infty} x_n$ is a maximal among all partial limits of $\{x_n\}$".

Assertion2. The upper limit $\overline{\lim}_{n\to\infty} x_n$ has the next property: for any positive $\varepsilon > 0$ there exist $k \in \mathrm{N}$, starting from which $\forall n > k \Rightarrow x_n < \overline{\lim}_{n\to\infty} x_n + \varepsilon$. And the lower limit $\underline{\lim}_{n\to\infty} x_n \equiv \inf \mathrm{P}$ has the similar property: for any positive $\varepsilon > 0$ there exist $k \in \mathrm{N}$, starting from which
$\forall n > k \Rightarrow x_n > \underline{\lim}_{n\to\infty} x_n - \varepsilon$.

Proof. Let's assume that it's not true, it means that there exist some "bad" $\bar{\varepsilon} > 0$ for which there is no any appropriate $k \in \mathrm{N}$, i.e., for any concrete $\bar{k} \in \mathrm{N}$ there exist at least one $\bar{n} > \bar{k}$ such that $x_{\bar{n}} \geq \overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$. So, for $k = 1$ there exist $n_1 > k = 1$ such that $x_{n_1} \geq \overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$. For $k = n_1$ there exist $n_2 > k = n_1$ such that $x_{n_2} \geq \overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$. For $k = n_2$ there exist $n_3 > k = n_2$ such that $x_{n_3} \geq \overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$ and etc. Then we have the subsequence $\{x_{n_k}\}$, and every it's term is not less than $\overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$. As the initial sequence $\{x_n\}$ is bounded, then $\{x_{n_k}\}$ is also bounded, then it has a convergent subsequence $\{x_{p_k}\}$ and $\{x_{p_k}\}$ is also a subsequence of $\{x_n\}$, and every it's term is not less than $\overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$, then it's limit is also not less than $\overline{\lim}_{n\to\infty} x_n + \bar{\varepsilon}$, then this limit is greater than $\overline{\lim}_{n\to\infty} x_n$, and we have a contradiction (because $\overline{\lim}_{n\to\infty} x_n$ is a maximal among all partial limits). This contradiction proves our assertion for an upper limit.
For $\underline{\lim}_{n\to\infty} x_n$ the proof is similar.

Assertion3. The sequence $\{x_n\}$ converges to some limit $a \Leftrightarrow \underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = a$.

Proof. $\Rightarrow$ Let $\{x_n\}$ converges, so $\lim_{n\to\infty} x_n = a$ then any it's subsequence goes to the same limit a, in particular, the sequences $\{x_{n_k}\}$ and $\{x_{p_k}\}$ (from the assertion1) both go to a, then $\inf \mathrm{P} = \sup \mathrm{P} = a \Leftrightarrow \underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = a$.

Conversely $\Leftarrow$ We have $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = a$. Let's fix an arbitrary positive $\varepsilon > 0$, according to the assertion2 there exist k such that $\forall n > k$ we have $x_n < \overline{\lim}_{n\to\infty} x_n + \varepsilon \Leftrightarrow x_n < a + \varepsilon$ and there exist m such that $\forall n > m$ we have $x_n > \underline{\lim}_{n\to\infty} x_n - \varepsilon \Leftrightarrow a - \varepsilon < x_n$. Let's fix the maximal $\bar{k} = \max(m, k)$, then starting from $\bar{k}$, $\forall n > \bar{k}$ we have $a - \varepsilon < x_n < a + \varepsilon \Rightarrow x_n \in \mathrm{O}_\varepsilon(a)$.

So, for any positive $\varepsilon > 0$, there exist the number $\bar{k}$, starting from which every term of the sequence $\{x_n\}$ belongs to $O_\varepsilon(a)$, it means that $\lim_{n\to\infty} x_n = a$.

Let's consider now the case of an unbounded sequence $\{x_n\}$.

Assertion4. If non-negative (in particular positive) sequence $\{x_n\} \| x_n \geq 0 \ \forall n$ is unbounded, then it has some subsequence $\{x_{n_k}\}$ which goes to $+\infty$. If non-positive (in particular negative) sequence $\{x_n\} \| x_n \leq 0 \ \forall n$ is unbounded, then it has some subsequence $\{x_{n_k}\}$ which goes to $-\infty$.

Proof. Let's fix any non-negative sequence $\{x_n\} \| x_n \geq 0 \ \forall n$. As it is unbounded, for any positive $M > 0$ there exist some number k such that $x_k > M$. Then for $M = 1$ there exist n_1 such that $x_{n_1} > 1$, for $M = 2$ there exist $n_2 > n_1$ such that $x_{n_2} > 2$. (if there is no such n_2, then all the terms of $\{x_n\}$ starting from the number n_1 are not greater than 2 and there are also n_1 first terms $x_1, x_2 .... x_{n_1}$ which can be any real numbers. Such sequence is bounded, and we have a contradiction). Then for $M = 3$ we can find $n_3 > n_2$ such that $x_{n_3} > 3$ and etc. As a result we have the subsequence $\{x_{n_k}\}$ such that $x_{n_k} > k \| \forall k \in N$, then $\{x_{n_k}\} \to +\infty$.

Let now we have a non-positive unbounded sequence $\{x_n\}$, then $\{-x_n\}$ is a non-negative unbounded sequence, and, as we showed above, it has some subsequence $\{-x_{n_k}\} \to +\infty$, then $\{x_{n_k}\} \to -\infty$ and $\{x_{nk}\}$ is a subsequence of the initial sequence $\{x_n\}$.

Consequence (from the assertion4). For any non-negative sequence $\{x_n\}$ only one of the next two cases is true:

[1] $\{x_n\}$ is bounded and therefore (according to the assertion1) the upper limit $\overline{\lim}_{n\to\infty} x_n$ is a concrete real number.

[2] $\{x_n\}$ is unbounded and therefore it has some subsequence $\{x_{n_k}\}$ which goes to $+\infty$.

In this case, by definition, we write $\overline{\lim}_{n\to\infty} x_n = +\infty$.

This consequence has a great importance for serieses and power serieses, both will be considered in the next 3-rd book.

We have proved enough for practical purposes.

Derivative

Def. $f(x)$ is defined on some neighborhood $\mathrm{O}_R(a)$ of a, then the function $\dfrac{f(x)-f(a)}{x-a}$ is defined on the deleted neighborhood $D_R(a)$, if the limit $\lim_{x\to a}\dfrac{f(x)-f(a)}{x-a}\equiv \mathrm{A}$ exists, then it is called a derivative of $f(x)$ at the point a, and we denote $f_x(a)\equiv \mathrm{A}$. And f is called differentiable at a in this case.

Comment. In many books almost immediately after a derivative is defined, the "geometrical meaning" of it is described like a slope of a tangent line. But there is never a definition, what is a tangent line actually is. There is the chapter in this book about curves, and there will be a normal definition of a tangent line, and then we will formulate normally the geometrical meaning of a derivative. Now a derivative is a concrete limit of a concrete function.

Assertion1. f is defined on $\mathrm{O}_R(a)$ and f is differentiable at a $\Leftrightarrow$ the next representation is possible on $\mathrm{O}_R(a)$: $f(x)-f(a)=\mathrm{A}\cdot(x-a)+\alpha(x)\cdot(x-a)\,||\,\forall x\in \mathrm{O}_R(a)$, where $\mathrm{A}=f_x(a)$ is a constant and $\alpha(x)$ is infinitely small when $x\to a$ and continuous at $x=a$ (i.e., $\lim_{x\to a}\alpha(x)=0=\alpha(a)$).

Proof. $\Rightarrow$Let f is differentiable at a, it means that the next limit exists $\lim_{x\to a}\dfrac{f(x)-f(a)}{x-a}=f_x(a)$. The function $\dfrac{f(x)-f(a)}{x-a}-f_x(a)\equiv //\,by\ def\,//\equiv\tilde{\alpha}(x)$ **[T]** is defined on $D_R(a)$ and the limit of the left part of **[T]** when $x\to a$ is equal to zero, then the same is true for the right part of **[T]**, then $\lim_{x\to a}\tilde{\alpha}(x)=0$. From **[T]** follows that $\forall x\in D_R(a)$ the next representation is true $f(x)-f(a)=f_x(a)\cdot(x-a)+\tilde{\alpha}(x)\cdot(x-a)$ **[E]**. And now we have the representation **[E]** which is almost what we need, but this equality is true on $D_R(a)$, and we can't even substitute the value $x=a$ in **[E]**, because $\tilde{\alpha}(x)$ isn't defined at the point $x=a$ (really, look at the definition **[T]** of this function). But remember that we are proving now that the needed representation **[E]** is possible, then in our power to find any appropriate function $\alpha(x)$ which satisfies to **[E]**. Let's do it. We define the new function $\alpha(x)$ on $\mathrm{O}_R(a):\forall x\in D_R(a)\Rightarrow\alpha(x)\equiv\tilde{\alpha}(x)$ and $\alpha(a)\equiv 0$. Then the representation **[E1]** $f(x)-f(a)=f_x(a)\cdot(x-a)+\alpha(x)\cdot(x-a)$ is true for any $x\in \mathrm{O}_R(a)$ and $\lim_{x\to a}\alpha(x)=0=\alpha(a)$. Everything is proved.

Conversely$\Leftarrow$ Let $f(x)-f(a)=\mathrm{A}\cdot(x-a)+\alpha(x)\cdot(x-a)\,||\,\forall x\in \mathrm{O}_R(a)$, then the same representation is true on $D_R(a)$, so $f(x)-f(a)=\mathrm{A}\cdot(x-a)+\alpha(x)\cdot(x-a)\,||\,\forall x\in D_R(a)$ let's divide all the sides by $(x-a)$, then $\dfrac{f(x)-f(a)}{x-a}=\mathrm{A}+\alpha(x)$ the right part has a limit A when $x\to a$, then the left part also has a limit A when $x\to a$, and we have $\lim_{x\to a}\dfrac{f(x)-f(a)}{x-a}\equiv\mathrm{A}$, then f is differentiable at a and $\mathrm{A}=f_x(a)$, everything is proved.

Consequence1. f is differentiable at $a \Rightarrow$ f is continuous at a.

Proof. Let's use the representation from above:
$f(x) = f(a) + f_x(a)\cdot(x-a) + \alpha(x)\cdot(x-a) \,||\, \forall x \in \mathrm{O}_R(a)$ **[V]** every summand on the right side has a limit when $x \to a$. So $\lim_{x\to a} f(a) = f(a)$, $\lim_{x\to a} f_x(a)\cdot(x-a) = 0$ (*because* $f_x(a) = const$), $\lim_{x\to a} \alpha(x)\cdot(x-a) = 0$, then the right part of **[V]** has a limit $f(a)$ when $x \to a$, then the left part of **[V]** has the same limit $f(a)$ when $x \to a$, i.e., $\lim_{x\to a} f(x) = f(a)$ and it means that f is continuous at a.

Assertion2. If f, g are differentiable at a. Then

[A] $f(x) + g(x)$ is differentiable at a, and it's derivative at a is a sum of derivatives $f_x(a) + g_x(a)$

[B] $f(x)\cdot g(x)$ is differentiable at a, and it's derivative at a is $f_x(a)\cdot g(a) + g_x(a)\cdot f(a)$.

[C] If $g(a) \neq 0$, then $f(x)/g(x)$ is differentiable at a, and it's derivative at a is
$$\frac{f_x(a)\cdot g(a) - g_x(a)\cdot f(a)}{[g(a)]^2}.$$

Proof. f, g are differentiable at a, so both limits $\lim_{x\to a} \frac{f(x) - f(a)}{x-a} \equiv f_x(a)$ and $\lim_{x\to a} \frac{g(x) - g(a)}{x-a} \equiv g_x(a)$ do exist. f is defined on some $\mathrm{O}_R(a)$ and g is defined on some $\mathrm{O}_P(a)$, let's fix the smaller of these neighborhoods $\mathrm{O}_R(a), \mathrm{O}_P(a)$ (let it be $\mathrm{O}_R(a)$) both functions f, g are defined on $\mathrm{O}_R(a)$, then $f+g$ and $f\cdot g$ are defined on $\mathrm{O}_R(a)$, and therefore we can raise a question about their derivatives at a.

If the next limit $\lim_{x\to a} \frac{(f(x)+g(x)) - (f(a)+g(a))}{x-a}$ exists, then, by definition, it is a derivative of $f(x)+g(x)$ at a. So, $\frac{(f(x)+g(x)) - (f(a)+g(a))}{x-a} = \frac{f(x)-f(a)}{x-a} + \frac{g(x)-g(a)}{x-a}$ the right part has a limit $f'(a) + g'(a)$ when $x \to a$, then the left part has the same limit $f'(a)+g'(a)$ when $x \to a$.

Next, if the next limit $\lim_{x\to a} \frac{f(x)\cdot g(x) - f(a)\cdot g(a)}{x-a}$ exists, then, by definition, it is a derivative of $f(x)\cdot g(x)$ at a. Let's rewrite: **[1]**
$$\frac{f(x)\cdot g(x) - f(a)\cdot g(a)}{x-a} = \frac{f(x)\cdot g(x) - f(x)\cdot g(a) + f(x)\cdot g(a) - f(a)\cdot g(a)}{x-a} =$$
$$= \frac{f(x)\cdot(g(x)-g(a)) + g(a)\cdot(f(x)-f(a))}{x-a} = f(x)\cdot\frac{(g(x)-g(a))}{x-a} + g(a)\cdot\frac{(f(x)-f(a))}{x-a}.$$
The right part has a limit $f_x(a)\cdot g(a) + g_x(a)\cdot f(a)$ when $x \to a$. Then the left part **[1]** has the same limit when $x \to a$ and **[B]** is proved.

In the case **[C]** as $g(a)\neq 0$ and g is continuous at a (because g is differentiable at a) we can fix the neighborhood $O_\delta(a)$ of a, on which $g(x)$ has the same sign as a value $g(a)$, then $g(x)\neq 0\ \forall x\in O_\delta(a)$. As $f(x)$ is defined on $O_R(a)$ and $g(x)$ is defined on $O_\delta(a)$ and $g(x)\neq 0\ \forall x\in O_\delta(a)$, we take the smaller of these neighborhoods (let it be $O_\delta(a)$).

Then $f(x)/g(x)$ is defined on $O_\delta(a)$ and we can raise a question about it's derivative at a.

If the limit $\lim_{x\to a}\dfrac{f(x)/g(x)-f(a)/g(a)}{x-a}$ exists, then, by definition, it is a derivative of $f(x)/g(x)$ at a. Let's rewrite:

[2] $$\frac{f(x)/g(x)-f(a)/g(a)}{x-a}=\frac{f(x)\cdot g(a)-g(x)\cdot f(a)}{x-a}\cdot\frac{1}{g(x)\cdot g(a)}=$$

$$=\frac{f(x)\cdot g(a)-f(a)\cdot g(a)+f(a)\cdot g(a)-g(x)\cdot f(a)}{x-a}\cdot\frac{1}{g(x)\cdot g(a)}=$$

$$=\left(g(a)\cdot\frac{f(x)-f(a)}{x-a}-f(a)\cdot\frac{g(x)-g(a)}{x-a}\right)\cdot\frac{1}{g(x)\cdot g(a)}$$ this expression obviously has a limit $\left(g(a)\cdot f_x(a)-f(a)\cdot g_x(a)\right)\cdot\dfrac{1}{g^2(a)}$ when $x\to a$, then **[2]** has the same limit when $x\to a$ and **[C]** is proved.

In addition: **[D]** f is differentiable at a and $\lambda\in R$ is a constant, then $\lambda\cdot f(x)$ is differentiable at a and it's derivative at a is $\lambda\cdot f_x(a)$.

Def. For any functions f,g and any constants λ,μ, the function $\lambda f+\mu g$ is called a linear combination of functions f,g.

Consequence2. If f,g are differentiable at a, then any linear combination of these functions $\lambda f+\mu g$ is differentiable at a and it's derivative at a is $\lambda\cdot f_x(a)+\mu\cdot g_x(a)$. In particular, for $\lambda=1,\ \mu=-1$ the difference $f-\mu$ is differentiable at a, and it's derivative at a is $f_x(a)-g_x(a)$.

Def. Next symbols can be used in order to denote a derivative of f at a point a:

$$f_x(a),\ \ f'(a),\ \ \frac{\partial f}{\partial x}(a),\ \ \frac{\partial f(x)}{\partial x}\Big|_{x=a}.$$

By using these designations, the previous result (assertion2+consequence2) can be written as:

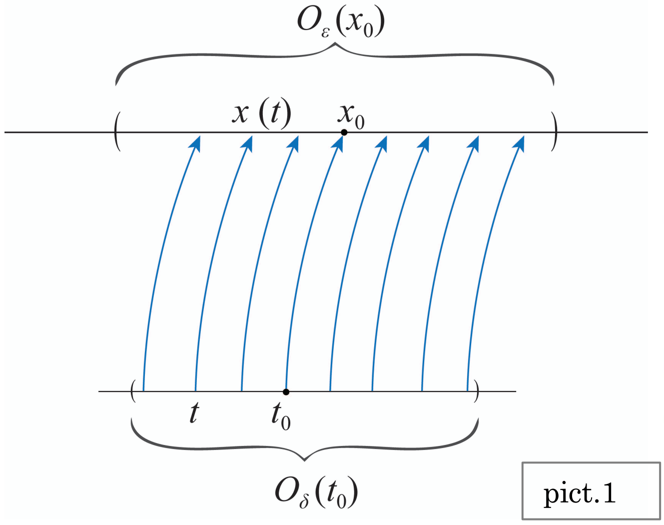
$$\frac{\partial(\lambda\cdot f(x)+\mu\cdot g(x))}{\partial x}\Big|_{x=a}=\lambda\cdot f_x(a)+\mu\cdot g_x(a)$$ and $$\frac{\partial(f(x)\cdot g(x))}{\partial x}\Big|_{x=a}=f(a)\cdot g_x(a)+f_x(a)\cdot g(a)$$

$$\frac{\partial(f(x)/g(x))}{\partial x}\Big|_{x=a}=\frac{f_x(a)\cdot g(a)-g_x(a)\cdot f(a)}{g^2(a)}.$$

Derivative of a composite function: $x = x(t)$ is differentiable at t_0 and $f = f(x)$ is differentiable at $x_0 = x(t_0)$. Then the composite function $f(x(t))$ is differentiable at t_0 and it's derivative is $f_t(x(t_0)) = f_x(x_0) \cdot x_t(t_0)$.

Proof. $x(t)$ is differentiable at t_0, then on some $\mathrm{O}_\delta(t_0)$ the next representation is true $x(t) - x(t_0) = \mathrm{A} \cdot (t - t_0) + \beta(t) \cdot (t - t_0) \,\|\, \forall t \in \mathrm{O}_\delta(t_0)$ **[3]**. And $f(x)$ is differentiable at x_0, then on some $\mathrm{O}_\varepsilon(x_0)$ the next representation is true $f(x) - f(x_0) = \mathrm{B} \cdot (x - x_0) + \alpha(x) \cdot (x - x_0) \,\|\, \forall x \in \mathrm{O}_\varepsilon(x_0)$ **[4]**.

The function $x(t)$ is continuous at t_0 (because it is differentiable at t_0), then we can squeeze (if necessary) the neighborhood $\mathrm{O}_\delta(t_0)$ **[pict1]**, in order to be: $\forall t \in \mathrm{O}_\delta(t_0)$ the value $x(t)$ belongs to $\mathrm{O}_\varepsilon(x(t_0)) \equiv \mathrm{O}_\varepsilon(x_0)$. Suppose we have already squeezed the neighborhood $\mathrm{O}_\delta(t_0)$.



pict.1

So, for any t from $\mathrm{O}_\delta(t_0)$ the representation **[3]** is true, and the value $x(t)$ belongs to $\mathrm{O}_\varepsilon(x_0)$, then for $x(t)$ the representation **[4]** is true, then $\forall t \in \mathrm{O}_\delta(t_0)$ we have $f(x(t)) - f(x_0) = \mathrm{B} \cdot (x(t) - x_0) + \alpha(x(t)) \cdot (x(t) - x_0)$ we can substitute here $x_0 = x(t_0)$, so $f(x(t)) - f(x(t_0)) = \mathrm{B} \cdot (x(t) - x(t_0)) + \alpha(x(t)) \cdot (x(t) - x(t_0)) \,\|\, \forall t \in \mathrm{O}_\delta(t_0)$ **[5]** and finally, we substitute the representation **[3]** of difference $x(t) - x(t_0)$ in **[5]**. Then $\forall t \in \mathrm{O}_\delta(t_0)$ we have $f(x(t)) - f(x(t_0)) = \mathrm{B} \cdot \big(\mathrm{A} \cdot (t - t_0) + \beta(t) \cdot (t - t_0)\big) + \alpha(x(t)) \cdot \big(\mathrm{A} \cdot (t - t_0) + \beta(t) \cdot (t - t_0)\big)$ it may look quite voluminous, but let's just regroup the summands on the right side:

$$f(x(t)) - f(x(t_0)) = \mathrm{B} \cdot \mathrm{A} \cdot (t - t_0) + \big[\mathrm{B} \cdot \beta(t) \cdot (t - t_0) + \alpha(x(t)) \cdot \big(\mathrm{A} \cdot (t - t_0) + \beta(t) \cdot (t - t_0)\big)\big] \Leftrightarrow$$
$$\Leftrightarrow f(x(t)) - f(x(t_0)) = [\mathrm{B} \cdot \mathrm{A}] \bullet (t - t_0) + \big[\mathrm{B} \cdot \beta(t) + \alpha(x(t)) \cdot \big(\mathrm{A} + \beta(t)\big)\big] \bullet (t - t_0) \ \textbf{[6]}.$$

Here $[\mathrm{B} \cdot \mathrm{A}]$ is a concrete constant (number), and the second expression in the square brackets $[\ldots.]$ is an infinitely small function when $t \to t_0$ and this expression is also a continuous at t_0 function. Really, let's consider: $\big[\mathrm{B} \cdot \beta(t) + \alpha(x(t)) \cdot \big(\mathrm{A} + \beta(t)\big)\big]$ **[R].** We obviously have here $\lim_{t \to t_0} \mathrm{B} \cdot \beta(t) = \mathrm{B} \cdot \lim_{t \to t_0} \beta(t) = 0$. Next, acc. to the theorem3 **[Continuity of a composite function]**: $x(t)$ is continuous at t_0 and $x(t_0) = x_0$, and $\alpha(x)$ is continuous at x_0, then the composite function $\alpha(x(t))$ is continuous at t_0, it means that $\lim_{t \to t_0} \alpha(x(t)) = \alpha(x(t_0)) = \alpha(x_0) = 0$ and finally, $\lim_{t \to t_0} \big(\mathrm{A} + \beta(t)\big) = \mathrm{A}$. From here follows that the limit of **[R]** when $t \to t_0$ is zero, then **[R]** is an infinitely small function when $t \to t_0$. All the separate functions from **[R]** are continuous at t_0, then **[R]** is continuous at t_0. Then we can write $\big[\mathrm{B} \cdot \beta(t) + \alpha(x(t)) \cdot \big(\mathrm{A} + \beta(t)\big)\big] \equiv \theta(t)$ where $\theta(t)$ is infinitely small when $t \to t_0$ and continuous at t_0. Then we can rewrite **[6]** as:

$f(x(t)) - f(x(t_0)) = [\mathrm{B}\cdot\mathrm{A}]\bullet(t-t_0) + \theta(t)\bullet(t-t_0)$ - it is the representation of difference $f(x(t)) - f(x(t_0))$ of a composite function $f(x(t))$ on the neighborhood $\mathrm{O}_\delta(t_0)$, and we see that this representation is exactly the same as we had in the assertion1.

Then $f(x(t))$ is differentiable at t_0 and it's derivative at t_0 is $\mathrm{B}\cdot\mathrm{A}$. From **[3]**,**[4]** we see that $\mathrm{B} = f_x(x_0)$ and $\mathrm{A} = x_t(t_0)$. So $f_t(x(t_0)) = f_x(x_0)\cdot x_t(t_0)$.

Def. $f(x)$ is defined on some half-interval $(a-R, a]$, then $\dfrac{f(x)-f(a)}{x-a}$ is defined on $(a-R, a)$.

If there exist the next **left limit** at a, $\lim_{x\to a-}\dfrac{f(x)-f(a)}{x-a} \equiv \mathrm{A}$, then A is called a left derivative of f at a, and we denote it $f'_-(a)$. In this case we say that f is left differentiable at a.

Def. $f(x)$ is defined on some half-interval $[a, a+R)$, then $\dfrac{f(x)-f(a)}{x-a}$ is defined on $(a, a+R)$.

If there exist the next **right limit** at a, $\lim_{x\to a+0}\dfrac{f(x)-f(a)}{x-a} \equiv \mathrm{A}$, then A is called a right derivative of f at a, and we denote it $f'_+(a)$. In this case we say that f is right differentiable at a.

For the left/right derivative it's easy to prove the assertions that are exactly similar to the assertion1,assertion2 and consequence1, consequence2 (formulate them). Even the theorem about a derivative of a composite function can be extended: the function $x(t)$ is left/right differentiable (or just differentiable) at t_0 and $f(x)$ is left/right differentiable (or just differentiable) at $x_0 = x(t_0)$. And $f(x)$ is defined on every value $x(t)$, then $f(x(t))$ differentiable at t_0 in exactly the same way (left/right/ordinary) as $x(t)$ is differentiable at t_0. And the formula $f_t(x(t_0)) = f_x(x_0)\cdot x_t(t_0)$ is true (there may be right/left/ordinary derivatives in this formula).

From the basic properties of function limits we have:
f is differentiable at $a \Leftrightarrow f$ is left differentiable at a and f is right differentiable at a, and left and right derivatives at a are equal $f'_-(a) = f'_+(a) = f'(a)$.

Def. $f(x)$ is defined on some neighborhood $\mathrm{O}_R(a)$. We say that $f(x)$ has a local maximum at a if $f(a) \ge f(x)\ \forall x \in D_R(a)$, in particular a strict local maximum if $f(a) > f(x)\ \forall x \in D_R(a)$. And we say that $f(x)$ has a local minimum at a if $f(a) \le f(x)\ \forall x \in D_R(a)$, in particular a strict local minimum if $f(a) < f(x)\ \forall x \in D_R(a)$. If f has a local minimum or a local maximum at a, then we say that f has a local extremum at a.

Lemma1. f has a local extremum at a and f is differentiable at a, then $f_x(a)=0$.

Proof. $f(x)$ is defined on $O_R(a)$. The next limit exists $\lim_{x\to a}\dfrac{f(x)-f(a)}{x-a}\equiv A$. Let $A>0$, then we can fix some deleted neighbourhood $D_\delta(a)$, such that $\forall x\in D_\delta(a)$ the value $\dfrac{f(x)-f(a)}{x-a}$ **[J]** is positive. Let $x>a$ in $D_\delta(a)$, then $x-a>0$ and in order **[J]** to be positive there must be $f(x)>f(a)$. And similarly, when $x<a$ in $D_\delta(a)$, there must be $f(x)<f(a)$ (so on the right of a we have $f(x)>f(a)$, and on the left of a we have $f(x)<f(a)$). Then there can't be any local extremum at a. And we have a contradiction. Then the assumption $A>0$ was false.
The similar contradiction will appear if we assume $A<0$. Then $A=0$.

Def. f is defined on $[a,b]$. We say that f is differentiable on $[a,b]$ if f is differentiable in the ordinary way at any point $x_0\in(a,b)$ and f is right differentiable at a and left differentiable at b.

Obviously, if f is differentiable on $[a,b]$, then f is continuous on $[a,b]$.

Roll's theorem. f is differentiable on $[a,b]$ and $f(a)=f(b)$. Then there exist $x_0\in(a,b)$ such that $f_x(x_0)=0$.

Proof. f is differentiable on $[a,b]\Rightarrow f$ is continuous on $[a,b]$. From the Weierstrass theorem follows that it reaches it's maximal and minimal values at some points $c,d\in[a,b]$.
If $c\in(a,b)$, then c is a point of local extremum of $f(x)$, then (lemma1) $f_x(c)=0$ and c is the point we need $c\equiv x_0$. And **if** $d\in(a,b)$ we get exactly the same result. **If** $c\notin(a,b)$ and $d\notin(a,b)$, then c,d are the end points of the segment $[a,b]$. As $f(a)=f(b)$, then $f(c)=f(d)$ and the maximal and minimal values of f on $[a,b]$ are equal, then f is a constant function on $[a,b]$, so $f\equiv const$ on $[a,b]$, for such function it's derivative is equal to zero at any point $x_0\in(a,b)$, so any point $x_0\in(a,b)$ is appropriate.

Cauchy formula. f and g are differentiable on $[a,b]$ and $g_x(x_0)\neq 0\ \forall x_0\in(a,b)$.

Then there exist $x_0\in(a,b)$ such that: $\dfrac{f(b)-f(a)}{g(b)-g(a)}=\dfrac{f_x(x_0)}{g_x(x_0)}$.

Proof. Let's define the auxiliary function $\varphi(x)=[f(b)-f(a)]\cdot g(x)-[g(b)-g(a)]\cdot f(x)$- it is a linear combination of differentiable on $[a,b]$ functions, therefore $\varphi(x)$ is differentiable on $[a,b]$.
As $\varphi(x)$ satisfies to the Roll's theorem, there exist $x_0\in(a,b)$ such that $\varphi_x(x_0)=0$.
Both $[f(b)-f(a)]$ and $[g(b)-g(a)]$ are constants, then the derivative $\varphi_x(x_0)$ is
$\varphi_x(x_0)=[f(b)-f(a)]\cdot g_x(x_0)-[g(b)-g(a)]\cdot f_x(x_0)$, and we have
$[f(b)-f(a)]\cdot g_x(x_0)-[g(b)-g(a)]\cdot f_x(x_0)=0$, now we can divide both sides by $g_x(x_0)\neq 0$, and we will get the Cauchy formula.

Consequence1 [the mean value theorem]. f is differentiable on $[a,b]$. Then there exist $x_0 \in (a,b)$ such that $f(b)-f(a)=f_x(x_0)\cdot(b-a)$.

Proof. Let's take the pair of functions $f(x), g(x)\equiv x$ on $[a,b]$, the Cauchy formula can be used here, obviously $g_x(x_0)=1$ for any $x_0 \in (a,b)$, then we have:

$$\frac{f(b)-f(a)}{b-a}=\frac{f_x(x_0)}{1} \Rightarrow f(b)-f(a)=f_x(x_0)\cdot(b-a).$$

Comment. It's very important to notice that both the Cauchy formula and the mean value theorem can be applied even in the case when $b<a$. So we shouldn't care and check all the time that b is greater than a. If f is differentiable on some segment with ends a,b, then the mean value theorem $f(b)-f(a)=f_x(x_0)\cdot(b-a)$ is true. Here x_0 is some point in the interval with ends a,b. And the same for the Cauchy formula (explain why).

Consequence2 from [the mean value theorem]. f is differentiable on $[a,b]$. If $f_x(x_0)=0$ for any $x_0 \in (a,b)$, then $f(x)\equiv const$ on $[a,b]$. If $f_x(x_0)>0$ for any $x_0 \in (a,b)$, then f is strictly increasing on $[a,b]$. If $f_x(x_0)<0$ for any $x_0 \in (a,b)$, then f is strictly decreasing on $[a,b]$.

Proof. Let's fix any $x_1<x_2 \,||\, x_1,x_2 \in [a,b]$. As f is differentiable on $[x_1,x_2]$ then (the mean value theorem) there exist $x_0 \in (x_1,x_2)$ such that $f(x_2)-f(x_1)=f_x(x_0)\cdot(x_2-x_1)$ **[V]**, where x_0 is some point from the interval (x_1,x_2). If $f_x(x_0)=0 \,||\, \forall x_0 \in (a,b)$, then $f(x_2)=f(x_1)$ and it is true for any $x_1,x_2 \in [a,b]$, then f is a constant function on $[a,b]$. Similarly, if $f_x(x_0)>0 \,||\, \forall x_0 \in (a,b)$, then the right part of **[V]** is positive, then $f(x_2)>f(x_1)$ and f is strictly increasing on $[a,b]$. And similarly for the other case.

Derivative of an inverse function. $y=f(x)$ is defined on $[a,b]$ and $f_x(x)>0$ on $[a,b]$. Then the inverse function $\varphi(y) \,||\, y \in [f(a),f(b)]$ is differentiable at any point $f(x_0)$ and $\varphi_y(f(x_0))=1/f_x(x_0)$.

Proof. As $f_x(x)>0$ on $[a,b]$, then (consequence2) $f(x)$ is strictly increasing on $[a,b]$ also f is continuous on $[a,b]$ (because f is differentiable on $[a,b]$), then f is one-to-one mapping $[a,b]\to[f(a),f(b)]$, and the inverse function $\varphi:[f(a),f(b)]\to[a,b]$ is defined (we showed it in the inverse function theorem), and $\varphi(y)$ is also strictly increasing on $[f(a),f(b)]$ and continuous on $[f(a),f(b)]$. Let's fix an arbitrary point $y_0 \in [f(a),f(b)]$, there exist the unique $x_0 \in [a,b]$ such that $y_0=f(x_0)$.

We are interested in the next limit $\lim_{y\to y_0 || y\in[f(a),f(b)]} \frac{\varphi(y)-\varphi(y_0)}{y-y_0}$ **[L]**.

Let's fix an arbitrary sequence $\{y_n\} \subset [f(a), f(b)] \,||\, y_n \neq y_0 \ \forall n \,||\, \{y_n\} \to y_0$. If we show that $\left\{\dfrac{\varphi(y_n)-\varphi(y_0)}{y_n - y_0}\right\} \to \dfrac{1}{f_x(x_0)}$, then (by definition) the limit **[L]** is equal to $\dfrac{1}{f_x(x_0)}$, and it is exactly what we need.

We consider the sequence $\left\{\dfrac{\varphi(y_n)-\varphi(y_0)}{y_n - y_0}\right\}$, every y_n can be uniquely represented as $y_n = f(x_n)$, then $\left\{\dfrac{\varphi(y_n)-\varphi(y_0)}{y_n - y_0}\right\} = \left\{\dfrac{\varphi(f(x_n))-\varphi(f(x_0))}{f(x_n)-f(x_0)}\right\} = \left\{\dfrac{x_n - x_0}{f(x_n)-f(x_0)}\right\}$ **[L1]**. Let's notice that $x_n \neq x_0 \ \forall n$ (because if some $x_k = x_0$, then $f(x_k) = f(x_0) \Leftrightarrow y_k = y_0$, which contradicts to the choice of the sequence $\{y_n\}$). Then **[L1]** can be rewritten as $\left\{\dfrac{1}{\dfrac{f(x_n)-f(x_0)}{x_n - x_0}}\right\}$ **[L2]**. What is a limit of the sequence **[L2]**? We have $\{y_n\} \to y_0$ and φ is continuous at y_0, then $\{\varphi(y_n)\} \to \varphi(y_0) \Leftrightarrow \{x_n\} \to x_0$.

As $f(x)$ is differentiable at x_0, then the limit $\lim_{n\to\infty} \dfrac{f(x)-f(x_0)}{x - x_0} = f_x(x_0) > 0$ exists, and for the sequence $\{x_n\} \to x_0$ we have $\left\{\dfrac{f(x_n)-f(x_0)}{x_n - x_0}\right\} \to f_x(x_0)$, then the sequence **[L2]** goes to $\dfrac{1}{f_x(x_0)}$, everything is proved.

Comment. The same theorem is true when $f_x(x) < 0$ on $[a,b]$ (the proof is similar).

Let $f(x)$ is defined on $\mathrm{O}_R(a)$, then $\dfrac{f(x)-f(a)}{x-a}$ is defined on $D_R(a)$, and if this function has a limit at a, then it is a derivative $f_x(a)$. In practice it's usually more convenient to calculate the other limit, let's denote $\Delta x \equiv x - a$, then $x = a + \Delta x$ and we have $\dfrac{f(x)-f(a)}{x-a} = \dfrac{f(a+\Delta x)-f(a)}{\Delta x}$. We understand intuitively that when x goes to a, the difference Δx goes to zero, and we can calculate the limit $\lim_{\Delta x\to 0} \dfrac{f(a+\Delta x)-f(a)}{\Delta x}$ in order to find $f_x(a)$. Let's formulate it as:

Exercise1. If $f(x)$ is defined on $\mathrm{O}_R(a)$, then $\dfrac{f(a+\Delta x)-f(a)}{\Delta x} \equiv \varphi(\Delta x)$ (a function of a variable Δx) is defined on $D_R(0)$. If the limit $\lim_{\Delta x\to 0} \dfrac{f(a+\Delta x)-f(a)}{\Delta x}$ exists, then the limit $\lim_{x\to a} \dfrac{f(x)-f(a)}{x-a}$ also exists and these limits are equal.

Exercise2. Prove that for any $a \in R$ we have $\frac{\partial \sin x}{\partial x}|_{x=a} = \cos a$.

Solution. Let's fix any $a \in R$. We are interested in the next limit $\lim_{\Delta x \to 0} \frac{\sin(a+\Delta x) - \sin a}{\Delta x}$ let's use the formula for difference of sines:

$$\frac{\sin(a+\Delta x)-\sin a}{\Delta x} = \frac{2\cos\frac{a+\Delta x+a}{2}\cdot\sin\frac{a+\Delta x-a}{2}}{\Delta x} = \frac{\cos\left(a+\frac{\Delta x}{2}\right)\cdot\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}} \quad \textbf{[U]}.$$

We have proved earlier that $\lim_{x\to 0}\frac{\sin x}{x} = 1$, from here follows that $\lim_{\Delta x \to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}} = 1$ and obviously $\lim_{\Delta x\to 0}\cos\left(a+\frac{\Delta x}{2}\right) = \cos a$. Then the limit **[U]** exists and it is equal to $\cos a$.

Exercise3. Prove that for any $a \in R$ we have $\frac{\partial \cos x}{\partial x}|_{x=a} = -\sin a$.

Consequence. As we have $tg(x) = \frac{\sin x}{\cos x}$, $ctg(x) = \frac{\cos x}{\sin x}$, then by using the formula $\frac{\partial(f(x)/g(x))}{\partial x}|_{x=a} = \frac{f_x(a)\cdot g(a) - g_x(a)\cdot f(a)}{g^2(a)}$ it's easy to derive that at any point a, where tgx is defined, we have: $\frac{\partial tgx}{\partial x}|_{x=a} = \frac{1}{\sin^2 a}$ and at any point a, where $ctgx$ is defined, we have $\frac{\partial ctgx}{\partial x}|_{x=a} = -\frac{1}{\cos^2 a}$.

Exercise4. $\arcsin y$ is defined on $[-1,1]$, prove that for any $y_0 \in (-1,1)$ we have $\frac{\partial \arcsin y}{\partial y}|_{y=y_0} = \frac{1}{\sqrt{1-(y_0)^2}}$.

Solution: Let's consider the function $\sin x$ on the interval $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$, for any $x_0 \in \left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ we have $\frac{\partial \sin x}{\partial x}|_{x=x_0} = \cos x_0 > 0$, then, according to [derivative of an inverse function], the inverse function $x = \arcsin y$ is differentiable on $(-1,1)$ and for any $y_0 \in (-1,1) \,||\, y_0 = \sin x_0$ we have

$$\frac{\partial \arcsin y}{\partial y}|_{y=y_0} = \frac{1}{\frac{\partial \sin x}{\partial x}|_{x=x_0}} = \frac{1}{\cos x_0} = [as\ \ \cos x_0 > 0] = \frac{1}{\sqrt{1-\sin^2 x_0}} = \frac{1}{\sqrt{1-(y_0)^2}}.$$

Exercise5. Show that for any $y_0 \in (-1,1)$ we have $\frac{\partial \arccos y}{\partial y}|_{y=y_0} = -\frac{1}{\sqrt{1-(y_0)^2}}$.

And also, for any $y_0 \in R$ we have $\frac{\partial arctgy}{\partial y}|_{y=y_0} = \frac{1}{1+(y_0)^2}$ and $\frac{\partial arcctgy}{\partial y}|_{y=y_0} = -\frac{1}{1+(y_0)^2}$.

At this moment we can't find derivatives of logarithmic and power functions.
At first we need to derive several auxiliary limits.

Theorem1. $\lim_{x\to 0}(1+x)^{1/x} = e$.

Proof. We already have: $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$ (it is a definition of e)

Let's show at first that the right limit $\lim_{x\to 0+}(1+x)^{1/x}$ is equal to e, and then that the left limit $\lim_{x\to 0-}(1+x)^{1/x}$ is also equal to e, then the ordinary limit $\lim_{x\to 0}(1+x)^{1/x}$ exists and equal to e.

Let's fix an arbitrary sequence $\{x_n\} \subset (0,+\infty) \| x_n \neq 0 \ \forall n \| \{x_n\} \to 0$ if we show that $\{(1+x_n)^{1/x_n}\} \to e$, then there must be $\lim_{x\to 0+}(1+x)^{1/x} = e$. Shortly speaking, $\{x_n\}$ is a positive sequence that goes to zero.

As $\{x_n\} \to 0$, then, starting from some number k, there must be: $\forall m > k$: $0 < x_m < 1$, then we can discard first k terms $x_1, x_2 .... x_k$ from $\{x_n\}$ which do not necessary belong to $(0,1)$. So, we can assume that from the very beginning we have the sequence $\{x_n\} \to 0$ such that $0 < x_n < 1 \| \forall n$, then we have $1/x_n > 1 \| \forall n$. For every real number $1/x_n$ there exist the pair of natural numbers $m_n - 1,\ m_n$ such that $m_n \le 1/x_n < m_n + 1$ **[S]**. We obviously have $\{1/x_n\} \to +\infty$, then from **[S]** the sequence of natural numbers $\{m_n + 1\}$ also goes to $+\infty$.

Then $\left\{\left(1+\frac{1}{m_n+1}\right)^{m_n+1}\right\}$ is a subsequence of $\left(1+\frac{1}{n}\right)^n$ and therefore $\left\{\left(1+\frac{1}{m_n+1}\right)^{m_n+1}\right\} \to e$ **[V1]**.

As $\{m_n + 1\} \to +\infty$, then also $\{m_n\} \to +\infty$ and therefore $\left\{\left(1+\frac{1}{m_n}\right)^{m_n}\right\} \to e$ **[V2]**.

It's easy to get the next estimation for $(1+x_n)^{m_n}$, by using **[S]** and the basic properties of power function: $\left(1+\frac{1}{m_n+1}\right)^{m_n} < (1+x_n)^{1/x_n} < \left(1+\frac{1}{m_n}\right)^{m_n+1}$. From **[V1]** and **[V2]** follows that the outer sequences here both go to e, then from the squeeze theorem for sequences follows that $(1+x_n)^{1/x_n}$ also goes to e.

Let's show now that $\lim_{x\to 0-}(1+x)^{1/x}=e$. We fix again an arbitrary sequence $\{x_n\}\subset(-\infty,0)\,||\,x_n\neq 0\ \ \forall n\,||\,\{x_n\}\to 0$ we need to show that $\{(1+x_n)^{1/x_n}\}$ goes to e.

Similarly, as we assumed earlier, we can assume now that $-1<x_n<0\,||\,\forall n$.

$$(1+x_n)^{1/x_n}=(1-|x_n|)^{-1/|x_n|}=\frac{1}{(1-|x_n|)^{1/|x_n|}}=\left(\frac{1}{1-|x_n|}\right)^{1/|x_n|}=\left(1+\frac{|x_n|}{1-|x_n|}\right)^{1/|x_n|}=\left(\left(1+\frac{|x_n|}{1-|x_n|}\right)^{\frac{1-|x_n|}{|x_n|}}\right)^{\frac{1}{1-|x_n|}}\quad \textbf{[Y]}$$

Let's denote now $y_n\equiv\dfrac{|x_n|}{1-|x_n|}$ - it is a positive sequence, as $\{|x_n|\}$ goes to zero, then $\{y_n\}$ also goes to zero, and we can rewrite **[Y]** as

$\left((1+y_n)^{1/y_n}\right)^{y_n+1}=(1+y_n)^{(y_n+1)/y_n}=(1+y_n)^{1+1/y_n}=(1+y_n)\cdot(1+y_n)^{1/y_n}$ **[Y1]**. We obviously have $(1+y_n)\xrightarrow[n\to+\infty]{}1$ and, as we proved above: $\{(1+y_n)^{1/y_n}\}\xrightarrow[n\to+\infty]{}e$ (because $\{y_n\}$ is a positive sequence that goes to zero), then **[Y1]** goes to e when $n\to\infty$. Everything is proved.

Theorem2: [A] $\lim_{x\to 0}\dfrac{\ln(1+x)}{x}=1$ and **[B]** For any $a>0$ we have $\lim_{x\to 0}\dfrac{a^x-1}{x}=\ln a$.

Proof. Let's prove **[A]**. The function $\dfrac{\ln(1+x)}{x}$ is defined on $(-1,0)\cup(0,+\infty)$, we fix an arbitrary sequence $\{x_n\}\subset(-1,0)\cup(0,+\infty)\,||\,x_n\neq 0\ \ \forall n\,||\,\{x_n\}\to 0$ from the **theorem1** follows that $\{(1+x_n)^{1/x_n}\}\to e$. The function $\ln(x)$ is continuous at e (because it is continuous on $(0,+\infty)$), then there must be $\{\ln((1+x_n)^{1/x_n})\}\to\ln e\Leftrightarrow\left\{\dfrac{\ln(1+x_n)}{x_n}\right\}\to 1$ and **[A]** is proved.

Let's prove **[B]**. We fix an arbitrary positive $a>0$, the function $\dfrac{a^x-1}{x}$ is defined on $(-\infty,0)\cup(0,+\infty)$, let's take an arbitrary sequence $\{x_n\}\to 0$ from this set. If $a=1$, then **[B]** is obviously true.

Let $a\neq 1$, we want to show that the sequence $\left\{\dfrac{a^{x_n}-1}{x_n}\right\}$ goes to 1 (it is exactly what we need).

In order to do it we need to represent this sequence in another form. As $a\neq 1$, the function $\log_a x$ is defined, and this function is one-to-one $(0,+\infty)\to(-\infty,+\infty)$. Then for any x_n there exist the unique positive real number, which we denote as $1+y_n$, such that $\log_a(1+y_n)=x_n$ **[R]**.

As $x_n=\log_a(1+y_n)\Leftrightarrow a^{x_n}=1+y_n\Rightarrow y_n=a^{x_n}-1$ **[J]**. As $\{x_n\}\to 0$, there must be $\{a^{x_n}\}\to 1$ (because a^x is continuous at $x=0$). And from **[J]** we see that $\{y_n\}\to 0\,||\,y_n>-1\ \ \forall n$.

Then $\dfrac{a^{x_n}-1}{x_n} == \dfrac{a^{\log_a(1+y_n)}-1}{\log_a(1+y_n)} = \dfrac{(1+y_n)-1}{\log_a(1+y_n)} = \dfrac{y_n}{\log_a(1+y_n)}$ **[Z].** Let's use here the formula $\log_a x = \dfrac{\log_b x}{\log_b a}$, (we change the base of the logarithm) $\log_a(1+y_n) = \dfrac{\ln(1+y_n)}{\ln a}$, then $\dfrac{y_n}{\log_a(1+y_n)} = \dfrac{y_n \cdot \ln a}{\ln(1+y_n)}$ **[Z]**. from **[A]** follows that $\left\{\dfrac{y_n}{\ln(1+y_n)}\right\} \to 1$. And therefore, the limit of the sequence **[Z]** is $\ln a$. Everything is proved.

Now we are ready to find the derivatives of logarithmic and power functions:

Theorem3. Let $a>0 \,||\, a \neq 1$ is a fixed positive number, then

[A] $\log_a x$ is differentiable at any point $x_0 \in (0,+\infty)$ and $\dfrac{\partial \log_a x}{\partial x}\Big|_{x=x_0} = \dfrac{1}{x_0 \ln a}$.

[B] Let $a>0 \,||\, a \neq 1$, then the power function a^x is differentiable at any point $x_0 \in (-\infty,+\infty)$ and $\dfrac{\partial a^x}{\partial x}\Big|_{x=x_0} = a^{x_0} \ln a$.

Proof. **[A]** Let's fix an arbitrary $x_0 \in (0,+\infty)$. We need to calculate the limit:

$$\lim_{\Delta x\to 0} \frac{\log_a(x_0+\Delta x)-\log_a x_0}{\Delta x} = \lim_{\Delta x\to 0} \frac{\log_a(x_0+\Delta x)/x_0}{\Delta x} = \lim_{\Delta x\to 0} \frac{\log_a\left(1+\dfrac{\Delta x}{x_0}\right)}{\Delta x} =$$

$$= \lim_{\Delta x\to 0} \frac{\log_a\left(1+\dfrac{\Delta x}{x_0}\right)}{\dfrac{\Delta x}{x_0}} \cdot \frac{1}{x_0} = \frac{1}{x_0} \cdot \lim_{\Delta x\to 0} \frac{\log_a\left(1+\dfrac{\Delta x}{x_0}\right)}{\dfrac{\Delta x}{x_0}} = \begin{bmatrix} \log_a x = \\ = \dfrac{\log_b x}{\log_b a} \end{bmatrix} = \frac{1}{x_0} \cdot \lim_{\Delta x\to 0} \frac{\ln\left(1+\dfrac{\Delta x}{x_0}\right)}{\ln a \cdot \dfrac{\Delta x}{x_0}} =$$

$$= \frac{1}{x_0 \cdot \ln a} \cdot \lim_{\Delta x\to 0} \frac{\ln\left(1+\dfrac{\Delta x}{x_0}\right)}{\dfrac{\Delta x}{x_0}} = [Theorem2\ [A]] = \frac{1}{x_0 \cdot \ln a}.$$

[B] Let's fix an arbitrary $x_0 \in (-\infty,+\infty)$. We need to calculate the next limit:

$$\lim_{\Delta x\to 0} \frac{a^{x_0+\Delta x}-a^{x_0}}{\Delta x} = \lim_{\Delta x\to 0} a^{x_0} \cdot \left(\frac{a^{\Delta x}-1}{\Delta x}\right) = a^{x_0} \cdot \lim_{\Delta x\to 0} \left(\frac{a^{\Delta x}-1}{\Delta x}\right) = [Theorem2\ [B]] = a^{x_0} \cdot \ln a.$$

Theorem4.

[A] The power function $x^n \,||\, n \in \mathrm{N},\ \ x \in (-\infty,+\infty)$ is differentiable at any point $x_0 \in (-\infty,+\infty)$ and $\dfrac{\partial x^n}{\partial x}\big|_{x=x_0} = n \cdot {x_0}^{n-1}$. For any negative integer number n the function x^n is differentiable at any point $x_0 \in (-\infty,0) \cup (0,+\infty)$ and we have the same formula $\dfrac{\partial x^n}{\partial x}\big|_{x=x_0} = n \cdot {x_0}^{n-1}$.

[B] The generalized power function $x^a \,||\, x > 0,\ \ a \in R$ (which is by definition $x^a \equiv e^{a\ln x}$) is differentiable at any point $x_0 \in (0,+\infty)$ and $\dfrac{\partial x^n}{\partial x}\big|_{x=x_0} = a \cdot {x_0}^{a-1}$.

Proof. **[A]** Let's fix any $n \in \mathrm{N}$ and any $x_0 \in (-\infty,+\infty)$. We are interested in the next expression:

$$\frac{(x_0+\Delta x)^n - {x_0}^n}{\Delta x} = \begin{bmatrix} binomial \\ theorem \end{bmatrix} = \frac{\sum\limits_{k=0}^{n} C_n^k {x_0}^k \Delta x^{n-k} - {x_0}^n}{\Delta x} =$$

$$= \frac{C_n^0 {x_0}^0 \Delta x^n + C_n^1 {x_0}^1 \Delta x^{n-1} + C_n^2 {x_0}^2 \Delta x^{n-2} + \ldots + C_n^{n-2} {x_0}^{n-2} \Delta x^2 + C_n^{n-1} {x_0}^{n-1} \Delta x^1 + C_n^n {x_0}^n \Delta x^0 - {x_0}^n}{\Delta x} =$$

$$= \frac{C_n^0 {x_0}^0 \Delta x^n + C_n^1 {x_0}^1 \Delta x^{n-1} + C_n^2 {x_0}^2 \Delta x^{n-2} + \ldots + C_n^{n-2} {x_0}^{n-2} \Delta x^2 + C_n^{n-1} {x_0}^{n-1} \Delta x^1}{\Delta x} =$$

$= C_n^0 {x_0}^0 \Delta x^{n-1} + C_n^1 {x_0}^1 \Delta x^{n-2} + C_n^2 {x_0}^2 \Delta x^{n-3} + \ldots + C_n^{n-2} {x_0}^{n-2} \Delta x^1 + C_n^{n-1} {x_0}^{n-1}$ - every summand here is a function of Δx and every summand (except the last one) has a limit 0 when $\Delta x \to 0$, and the last summand has a limit $C_n^{n-1} {x_0}^{n-1}$ when $\Delta x \to 0$. Then all the sum has a limit $C_n^{n-1} {x_0}^{n-1}$ when $\Delta x \to 0$. Then $\lim_{\Delta x \to 0} \dfrac{(x_0+\Delta x)^n - {x_0}^n}{\Delta x} = C_n^{n-1} {x_0}^{n-1} = n {x_0}^{n-1}$.

Let now n is a negative integer number, then $x^n = \dfrac{1}{x^{|n|}}$. Let's fix an arbitrary $x_0 \in (-\infty,0) \cup (0,+\infty)$ and now we will use the formula $(f/g)_x(a) = \dfrac{f_x(a) \cdot g(a) - g_x(a) \cdot f(a)}{g^2(a)}$.

We can use it, because functions 1 and $x^{|n|},\ |n| \in \mathrm{N}$ are differentiable at any point $x_0 \in (-\infty,0) \cup (0,+\infty)$.

So $\left(\dfrac{1}{x^{|n|}}\right)_x (x_0) = \dfrac{0 \cdot {x_0}^{|n|} - |n| \cdot {x_0}^{|n|-1} \cdot 1}{{x_0}^{2|n|}} = \dfrac{-|n|}{{x_0}^{|n|+1}} = -|n| \cdot x^{-(|n|+1)} = -|n| \cdot x^{-|n|-1} = n \cdot x^{n-1}$.

[B] Let's consider $x^a \equiv e^{a\ln x} \,||\, x > 0,\ \ a \in R$. We fix any point $x_0 \in (0,+\infty)$. The function $e^{a\ln x}$ is a composite function $\varphi(g(x))$, where $\varphi(y) \equiv e^y,\ \ g(x) = a\ln x$. And $g(x)$ is differentiable at any

point $x_0 \in (0,+\infty)$ (we showed above that any logarithmic function is differentiable) and $\varphi(y)$ is differentiable at any point $g(x_0)$ (because the power function e^x is differentiable everywhere).

Then $\frac{\partial\varphi(g(x))}{\partial x}|_{x=x_0} = \frac{\partial\varphi(y)}{\partial y}|_{y=g(x_0)} \cdot \frac{\partial g(x)}{\partial x}|_{x=x_0}$, so we have

$$\frac{\partial e^{a\ln x}}{\partial x}|_{x=x_0} = \frac{\partial e^y}{\partial y}|_{y=a\cdot\ln x_0} \cdot \frac{\partial(a\cdot\ln x)}{\partial x}|_{x=x_0} = e^{a\cdot\ln x_0}\cdot a\cdot\frac{1}{x_0} = \left(e^{\ln x_0}\right)^a\cdot a\cdot\frac{1}{x_0} = x_0{}^a\cdot a\cdot\frac{1}{x_0} = a\cdot x_0{}^{a-1}.$$

Everything is proved.

Exercise6. Find the derivative of $f(x) = (x-a)^n \,||\, n\in\mathrm{N}$ at a point $x_0 \in R$.

Solution. This function can be considered as a composite function $\varphi(g(x))$, where $\varphi(y) = y^n$ and $g(x) = x - a$. As $g(x)$ is differentiable at any point x_0 and $\varphi(y)$ is differentiable at any point $y = \varphi(x_0)$, then $\varphi(g(x))$ is differentiable at any point x_0 and

$$\frac{\partial\varphi(g(x))}{\partial x}|_{x=x_0} = \frac{\partial\varphi(y)}{\partial y}|_{y=g(x_0)} \cdot \frac{\partial g(x)}{\partial x}|_{x=x_0}.$$

Then $\frac{\partial((x-a)^n)}{\partial x}|_{x=x_0} = \frac{\partial y^n}{\partial y}|_{y=(x_0-a)} \cdot \frac{\partial(x-a)}{\partial x}|_{x=x_0} = n\cdot(x_0-a)^{n-1}\cdot 1 = n\cdot(x_0-a)^{n-1}$.

Derivative examples

We will not consider the simpliest derivative-examples in this book, there is a huge amount of books and internet resources where it is in abundance. Let's take more advanced examples.

For Reader's practice:

[1] Find the derivative of $f(x) = \sin x^{\sin x^{\sin x^{\sin x^{\cdots}}}}$ at $x_0 = \pi/4$.

Solution. We have here $f(x) = \sin x^{f(x)} \Rightarrow \ln f(x) = f(x)\cdot\ln(\sin x)$ let's differentiate both sides with respect to x (when we do so, we need to take into account the rules of differentiation of a composite function).

$$(\ln f(x))' = (f(x)\cdot\ln(\sin x))' \Leftrightarrow \frac{1}{f(x)}\cdot f'(x) = f'(x)\cdot\ln(\sin x) + f(x)\cdot\frac{1}{\sin x}\cdot\cos x \Leftrightarrow$$

$$f'(x)\cdot\left(\frac{1}{f(x)} - \ln(\sin x)\right) = f(x)\cdot ctgx \Rightarrow f'(x) = \frac{f^2(x)\cdot ctgx}{1 - f(x)\cdot\ln(\sin x)}.$$

We have obtained the explicit expression for $f'(x)$, now we can substitute $x_0 = \pi/4$, then

$f'(\pi/4) = \frac{(f(\pi/4))^2\cdot ctg(\pi/4)}{1 - f(\pi/4)\cdot\ln(\sin(\pi/4))} \Leftrightarrow f'(\pi/4) = \frac{(f(\pi/4))^2}{1 - f(\pi/4)\cdot\ln(1/\sqrt{2})}$, we see that here appears the value $f(\pi/4)$, let's find it: $f(\pi/4) = \sin\pi/4^{\sin\pi/4^{\sin\pi/4^{\sin\pi/4^{\cdots}}}} = 1/\sqrt{2}^{1/\sqrt{2}^{1/\sqrt{2}^{1/\sqrt{2}^{\cdots}}}} \equiv a$.

We have here

$$a^{1/\sqrt{2}} = a \Rightarrow 1 = a^{1-1/\sqrt{2}} \Rightarrow \ln(1) = \ln(a^{1-1/\sqrt{2}}) \Leftrightarrow 0 = (1-1/\sqrt{2})\cdot \ln a \Rightarrow \ln a = 0 \Rightarrow a = e^0 = 1,$$

then we have $f(\pi/4) = 1$, then $f'(\pi/4) = \dfrac{1^2}{1-1\cdot\ln(1/\sqrt{2})} = \dfrac{1}{1-\ln(1/\sqrt{2})}$.

[2] Find the derivative of $f(x) = \ln x^{\ln x^{\ln x^{\ln x^{\cdots}}}}$ at $x_0 = e$.

Solution. Obviously $f(x) = \ln x^{f(x)} \Rightarrow \ln f(x) = f(x)\cdot\ln(\ln x)$ and we differentiate both sides with respect to x, so

$$\frac{1}{f(x)}\cdot f'(x) = f'(x)\cdot\ln(\ln x) + f(x)\cdot\frac{1}{\ln x}\cdot\frac{1}{x} \Rightarrow f'(x)\cdot\left(\frac{1}{f(x)} - \ln(\ln x)\right) = \frac{f(x)}{x\cdot\ln x} \Rightarrow$$

$\Rightarrow f'(x) = \dfrac{f^2(x)}{x\cdot\ln x\cdot(1-f(x)\cdot\ln(\ln x))}$. Let's substitute now $x = e$, $f'(e) = \dfrac{f^2(e)}{x\cdot\ln e\cdot(1-f(e)\cdot\ln(\ln e))}$

we obviously have $f(e) = 1$, then $f'(e) = \dfrac{1}{e\cdot 1\cdot(1-1\cdot\ln 1)} = \dfrac{1}{e}$.

[3] Find all the functions $f(x)$, that are defined on R, such that $|f(x)-f(y)| \le (x-y)^2$ $\forall x,y\in R$.

Answer: there must be $f(x)\equiv const$ on R. **Hint**. Show that $f'(x_0) = 0$ for any $x_0\in R$.

[4] $f(x), g(x), h(x)$ are differentiable on $[a,b]$. Show that there exist $c\in(a,b)$ such that

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(c) & g'(c) & h'(c) \end{vmatrix} = 0.$$

Solution. Let's consider the auxiliary function $\varphi(x) \equiv \begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f(x) & g(x) & h(x) \end{vmatrix}$, this function is obviously differentiable on $[a,b]$ and obviously $\varphi(a) = \varphi(b) = 0$, then there exist $c\in(a,b)$ such that $\varphi'(c) = 0$. And it's easy to understand (as $f(a), f(b), g(a)\ldots.$ and etc. are constants) that for any $x\in[a,b]$ we have $\varphi'(x) \equiv \begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(x) & g'(x) & h'(x) \end{vmatrix}$, we already know that $\varphi'(c) = 0$ - it is exactly what we need.

[5] Let $f(x) = x + \cfrac{1}{2x + \cfrac{1}{2x + \cfrac{1}{2x + \ldots.}}}$, find $f(100)$ and $f'(100)$.

Answer: $f(100) = \sqrt{10001}$ and $f'(100) = 100/\sqrt{10001}$.

Hint. In such cases we always need to create some simple auxiliary equality for $f(x)$ (as above).

[6] $f(x)$ is differentiable everywhere (on R) and it's derivative $f'(x)$ is continuous everywhere (such function f is called a "smooth function"). We have $f(0)=6$ and $f'(4)=5$.

Calculate $\lim_{x\to 2}\dfrac{f(4)-f(x^2)}{2-x}$. **Answer**: 20.

Solution. According to the mean value theorem, we can write $f(4)-f(x^2)=f'(c(x^2))\cdot(4-x^2)$ here $c(x)$ is some point inside the interval with ends $4, x^2$, we denote it in such way because for every value x there is a concrete value $c\equiv c(x)$. Then

$$\lim_{x\to 2}\frac{f(4)-f(x^2)}{2-x}=\lim_{x\to 2}\frac{f'(c(x))\cdot(4-x^2)}{2-x}=\lim_{x\to 2} f'(c(x))\cdot(2+x)\ \textbf{[L]}.$$

Obviously $\lim_{x\to 2}(2+x)=4$. But what is the value of $\lim_{x\to 2} f'(c(x))$? The simple explanation is: when $x\to 2\Rightarrow x^2\to 4$, it forces $c(x)\to 4$ and then $f'(c(x))\to f'(4)=5$ (initial condition), so $\lim_{x\to 2} f'(c(x))=4$ and the limit **[L]** is 20.

Notice. We got the equality $\lim_{x\to 2} f'(c(x))=4$, the explanation from above is quite convenient, but not rigorous. Let's give a normal explanation. We consider $f'(c(x))$ as a composite function, it is constructed from the functions $f'(y)\,||\,y=c(x)$. The value $c(x)$ is a number inside the interval with ends $4, x^2$. It's easy to understand that $\lim_{x\to 2} c(x)=4$ (just because $c(x)$ is always between 4 and x^2) and $z=f'(y)$ is continuous at $y=4$ (because it is continuous everywhere). Then (Improvement of Theorem3 [Limit of a composite function]) $\lim_{x\to 2} f'(c(x))=f'(4)=5$.

[7] $f(x)$ is differentiable at $x=a$ and $f'(a)=1/4$. Calculate $\lim_{h\to 0}\dfrac{f(a+2h^2)-f(a-2h^2)}{h^2}$.

Hint. Use the mean value theorem. **Answer**: 1.

[8] $f(x)$ is such function that $2^x+2^{f(x)}=2^{x+f(x)}$ find $f'(8)$.

Solution: $2^x+2^{f(x)}=2^{x+f(x)}\Leftrightarrow 2^{f(x)}(2^{x-f(x)}+1)=2^{f(x)}\cdot 2^x\Rightarrow 2^{x-f(x)}+1=2^x\Rightarrow$

$\Rightarrow 2^{x-f(x)}=2^x-1\Rightarrow \log_2 2^{x-f(x)}=\log_2(2^x-1)\Rightarrow x-f(x)=\log_2(2^x-1)\Rightarrow f(x)=x-\log_2(2^x-1)$

we got the explicit formula for $f(x)$, and now we can use the standard rules to find it's derivative:

$$f'(x)=1-(\log_2(2^x-1))'\cdot(2^x-1)',\ \text{so}\ f'(x)=1-\frac{1}{(2^x-1)\cdot\ln 2}\cdot(2^x\cdot\ln 2)=1-\frac{2^x}{2^x-1}=\frac{-1}{2^x-1},$$

then $f'(8)=\dfrac{-1}{2^8-1}=\dfrac{-1}{255}$.

[9] $f(x)$ is such function that $f(x)=\sqrt{x+\sqrt{x+\sqrt{x+.....}}}$, find $f'(2018)$. **Answer**: $f'(2018)=\dfrac{2}{\sqrt{8073}-1}$

[10] Functions g and f are both differentiable on $[a,b)$, here $b\in R$, or $b=+\infty$.
And $g(a)=f(a)$ and $g_x(x)<f_x(x)$ everywhere on (a,b). Show that $g(x)<f(x)$ everywhere on (a,b). **Hint**. Consider the auxiliary function $\varphi(x)\equiv f(x)-g(x)$.

Taylor's theorem

Higher Order Derivatives. $f(x)$ is defined on some neighborhood $\mathrm{O}_R(a)$ of a. Suppose that f is differentiable at any point $x_0 \in \mathrm{O}_R(a)$. Then for any $x_0 \in \mathrm{O}_R(a)$ the number $f_x(x_0)$ is defined and therefore $f_x(x)$ is an independent function, which is defined on $\mathrm{O}_R(a)$. And we can raise a question about it's derivative at a. If $f_x(x)$ is differentiable at a, i.e., the limit $\lim_{x\to a}\dfrac{f_x(x)-f_x(a)}{x-a}$ exists, then this limit is called the second derivative of f at a, and we denote it $f_{xx}(a) \equiv \lim_{x\to a}\dfrac{f_x(x)-f_x(a)}{x-a}$. Similarly, for any point $x_0 \in \mathrm{O}_R(a)$ we can raise a question about a derivative of $f_x(x)$ at x_0, i.e., the question about the existence of the limit $\lim_{x\to x_0}\dfrac{f_x(x)-f_x(x_0)}{x-x_0}$, if this limit exists, we denote it $f_{xx}(x_0)$ and it is called the second derivative of f at x_0.
If $f_x(x)$ is differentiable at any point $x_0 \in \mathrm{O}_R(a)$, then $f_{xx}(x)$ is an independent function on $\mathrm{O}_R(a)$ and we can raise a question about it's derivative at any point $x_0 \in \mathrm{O}_R(a)$, if such derivative exists, we denote it $f_{xxx}(x_0)$ and it is called the third derivative of f at x_0 and etc.

Def. An ***n-th*** derivative of f at a can be denoted as $f^n(x)$.

Assertion1. Let $f(x)$ is differentiable on $\mathrm{O}_R(a)$ and $f_x(a)=0$. And the second derivative $f_{xx}(a)$ exists.
[A] If $f_{xx}(a)>0$, then f has a strict local minimum at a.
[B] If $f_{xx}(a)<0$, then f has a strict local maximum at a.

Proof. By definition $f_{xx}(a)=\lim_{x\to a}\dfrac{f_x(x)-f_x(a)}{x-a}=\lim_{x\to a}\dfrac{f_x(x)}{x-a}$. If **[A]** $f_{xx}(a)>0$, then $\lim_{x\to a}\dfrac{f_x(x)}{x-a}>0$, then there exist the deleted neighborhood $D_\delta(a)\subset \mathrm{O}_R(a)$, on which $\dfrac{f_x(x)}{x-a}>0$.
So for any $x\in D_\delta(a)=(a-\delta,a)\cup(a,a+\delta)$ we have $\dfrac{f_x(x)}{x-a}>0$. If $x\in(a-\delta,a)$, then $x-a<0$ and there must be $f_x(x)<0$. So, for any $x\in(a-\delta,a)$ we have $f_x(x)<0$, then $f(x)$ is strictly decreasing on $(a-\delta,a]$. If $x\in(a,a+\delta)$, then $x-a>0$ and there must be $f_x(x)>0$.
So, for any $x\in(a,a+\delta)$ we have $f_x(x)>0$, then $f(x)$ is strictly increasing on $[a,a+\delta)$.
So f is strictly decreasing on $(a-\delta,a]$ and f is strictly increasing on $[a,a+\delta)$.
Then a is a point of a strict local minimum of f. And there is a similar proof for **[B]**.

Def. $f(x)$ is defined on $O_R(a)$ and it is n times differentiable at a.

The function $\mathrm{T}_n(x)\equiv f(a)+\dfrac{f^1(a)}{1!}\cdot(x-a)^1+\dfrac{f^2(a)}{2!}\cdot(x-a)^2+\ldots+\dfrac{f^n(a)}{n!}\cdot(x-a)^n$ is called a Taylor polynomial of $n-th$ degree for $f(x)$ with center at a.

Auxiliary1. $f(a)=\mathrm{T}_n(a)$ and $f_x(a)=(\mathrm{T}_n)_x(a)$, $f_{xx}(a)=(\mathrm{T}_n)_{xx}(a)$ ..... $f^n(a)=(\mathrm{T}_n)^n(a)$. (it's very easy to check these equalities).

Auxiliary2. $\varphi(x)$ and $g(x)$ are defined on $\mathrm{O}_R(a)$ and both these functions are $(n+1)$ times differentiable on $O_R(a)$ and also:

[A] $\varphi(a)=g(a)=0$, $\varphi_x(a)=g_x(a)=0$ .... $\varphi^n(a)=g^n(a)=0$,

[B] $g(x)\neq 0$ for any $x\in D_R(a)$ and for any $k\in\mathrm{N}$ we have $g^k(x)\neq 0 \,||\, \forall x\in D_R(a)$.

Then: for any $x\in D_R(a)$ there exist the point c from the interval with ends x and a such that:

$$\frac{\varphi(x)}{g(x)}=\frac{\varphi^{n+1}(c)}{g^{n+1}(c)}.$$

Proof. We fix any $x\in D_R(a)=(a-R,a)\cup(a,a+R)$, let $x\in(a,a+R)$.

Let's apply the Cauchy formula for $\varphi(x)$ and $g(x)$ on $[a,x]$, then we have

$\dfrac{\varphi(x)-\varphi(a)}{g(x)-g(a)}=\dfrac{\varphi_x(c_1)}{g_x(c_1)} \,||\, a<c_1<x$, according to **[A]**, we have $\varphi(a)=g(a)=0$, then we get

$\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_x(c_1)}{g_x(c_1)} \,||\, a<c_1<x$. According to **[A]**, we can rewrite it $\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_x(c_1)-\varphi_x(a)}{g_x(c_1)-g_x(a)} \,||\, a<c_1<x$.

Now we can apply the Cauchy formula for $\varphi_x(x)$ and $g_x(x)$ on $[a,c_1]$, we will get

$\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_x(c_1)-\varphi_x(a)}{g_x(c_1)-g_x(a)}=\dfrac{\varphi_{xx}(c_2)}{g_{xx}(c_2)} \,||\, a<c_2<c_1<x$, so now we have $\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_{xx}(c_2)}{g_{xx}(c_2)} \,||\, a<c_2<c_1<x$.

According to **[A]**, we can rewrite it: $\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_{xx}(c_2)-\varphi_{xx}(a)}{g_{xx}(c_2)-\varphi_{xx}(a)} \,||\, a<c_2<c_1<x$ and now we apply the Cauchy formula for $\varphi_{xx}(x)$ and $g_{xx}(x)$ on $[a,c_2]$.

Then we have $\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi_{xxx}(c_3)}{g_{xxx}(c_3)} \,||\, a<c_3<c_2<c_1<x$ and etc. Eventually we will get

$\dfrac{\varphi(x)}{g(x)}=\dfrac{\varphi^{n+1}(c)}{g^{n+1}(c)} \,||\, a<c<x$. And there is a similar proof for any fixed $x\in(a-R,a)$.

Taylor's theorem. f is defined on $O_R(a)$ and $(n+1)$ times differentiable on $O_R(a)$.

Then for any $x\in D_R(a)$ the next formula is true: $f(x)=\mathrm{T}_n(x)+\dfrac{f^{n+1}(c)}{(n+1)!}\cdot(x-a)^{n+1}$ **[V]** where c is some point from the interval with ends x and a, and $\mathrm{T}_n(x)$ is a Taylor polynomial (def above).

Proof. Let's fix an arbitrary $x\in D_R(a)$. From auxiliary1 follows that for $\varphi(x)\equiv f(x)-\mathrm{T}_n(x)$ we have $\varphi(a)=0$, $\varphi_x(a)=0$, $\varphi_{xx}(a)=0$.... $\varphi^n(a)=0$. Let's take the function $g(x)\equiv(x-a)^{n+1}$, then the pair $\varphi(x),g(x)$ satisfies to the auxiliary2. Then there exist c from the interval with ends

a, x such that:

$$\frac{\varphi(x)}{g(x)} = \frac{\varphi^{n+1}(c)}{g^{n+1}(c)} \Leftrightarrow \frac{f(x)-\mathrm{T}_n(x)}{(x-a)^{n+1}} = \frac{f^{n+1}(c)-\mathrm{T}_n^{\ n+1}(c)}{g^{n+1}(c)} \Leftrightarrow \left[\begin{matrix} as\ \mathrm{T}_n^{\ n+1}(x)\equiv 0 \\ and\ g^{n+1}(x)\equiv (n+1)! \end{matrix}\right] \Leftrightarrow \frac{f(x)-\mathrm{T}_n(x)}{(x-a)^{n+1}} = \frac{f^{n+1}(c)}{(n+1)!} \Rightarrow$$

$\Rightarrow f(x)-\mathrm{T}_n(x) = \dfrac{f^{n+1}(c)}{(n+1)!}\cdot(x-a)^{n+1}$ - it is exactly what we need.

Notice that any value $x \in D_R(a)$ defines one point (the number) c (which lies between x and a), for which **[V]** is true. Then $c \equiv c(x)$ is a function which is defined on $D_R(a)$ and obviously $\lim_{x\to a} c(x) = a$ (just because $c(x)$ is always between x and a).

Def. The function $\theta(x) \equiv \dfrac{f^{n+1}(c)}{(n+1)!}\cdot(x-a)^{n+1}$ is called a remainder.

Assertion1. If $f^{n+1}(x)$ is continuous at a, then the remainder $\theta(x)$ can be represented as $\theta(x) \equiv \tilde{\alpha}(x)\cdot(x-a)^n$, where $\alpha(x)$ is a function on $D_R(a)$, which is infinitely small when $x \to a$.

Proof. By definition, the remainder is:

$$\theta(x) \equiv \frac{f^{n+1}(c(x))}{(n+1)!}\cdot(x-a)^{n+1} = \left[\frac{f^{n+1}(c(x))}{(n+1)!}\cdot(x-a)\right]\cdot(x-a)^n.$$ The function

$\tilde{\alpha}(x) \equiv \dfrac{f^{n+1}(c(x))}{(n+1)!}\cdot(x-a)$ is defined for every $x \in D_R(a)$. We need to show that this function goes to zero when $x \to a$, i.e., to show that $\lim_{x\to a}\tilde{\alpha}(x) = 0$. Let's fix an arbitrary sequence $\{x_n\} \subset \mathrm{O}_R(a) \,\|\, x_n \neq a\ \forall n \,\|\, \{x_n\} \to a$, every x_n defines one point $c_n = c_n(x_n)$ from the interval with ends x_n, a, then we have the sequence $\{c_n\} \equiv \{c_n(x_n)\} \subset \mathrm{O}_R(a) \,\|\, c_n \neq a\ \forall n \,\|\, \{c_n\} \to a$.

As $f^{n+1}(x)$ is continuous at a, there must be $\{f^{n+1}(c_n)\} \to f^{n+1}(a)$ or the same $\{f^{n+1}(c_n(x_n))\} \to f^{n+1}(a)$. Then the sequence $\{\tilde{\alpha}(x_n)\} = \left\{\dfrac{f^{n+1}(c(x_n))}{(n+1)!}\cdot(x_n-a)\right\}$ goes to zero 0, then $\lim_{x\to a}\tilde{\alpha}(x) = 0$. So $\tilde{\alpha}(x)$ is infinitely small when $x \to a$, everything is proved.

Consequence1. By definition of the symbol o, we can write $\theta(x)$ as $\theta(x) \equiv o((x-a)^n)$, which is equivalent to $\theta(x) \equiv \tilde{\alpha}(x)\cdot(x-a)^n$.

Let's sum up:

[1-st result] If $f(x)$ is defined on $O_R(a)$ and $(n+1)$ times differentiable on $O_R(a)$, then for any $x \in D_R(a)$ there exist some c from the interval with the ends x, a such that:

$$f(x) \equiv \left[f(a) + \frac{f^1(a)}{1!}\cdot(x-a)^1 + \frac{f^2(a)}{2!}\cdot(x-a)^2 + \ldots + \frac{f^n(a)}{n!}\cdot(x-a)^n\right] + \frac{f^{n+1}(c)}{(n+1)!}\cdot(x-a)^{n+1}$$ **[T]**.

[2-nd result] If in addition $f^{n+1}(x)$ is continuous at a, then the next representation is true on $D_R(a)$:

$$f(x)=\left[f(a)+\frac{f^1(a)}{1!}\cdot(x-a)^1+\frac{f^2(a)}{2!}\cdot(x-a)^2+\ldots+\frac{f^n(a)}{n!}\cdot(x-a)^n\right]+\tilde{\alpha}(x)\cdot(x-a)^n \text{ [M]},$$ where

$\lim_{x\to a}\tilde{\alpha}(x)=0$. Let's define now the new function $\alpha(x)$ on $O_R(a)$: $\forall x\in D_R(a)\Rightarrow\alpha(x)\equiv\tilde{\alpha}(x)$ and $\alpha(a)\equiv //\, by\ def\, //\equiv 0$, then for such function we have $\lim_{x\to a}\alpha(x)=0=\alpha(0)$, i.e., $\alpha(x)$ is infinitely small when $x\to a$ and also continuous at a. And for any $x\in O_R(a)$ the next representation is true:

$$f(x)=\left[f(a)+\frac{f^1(a)}{1!}\cdot(x-a)^1+\frac{f^2(a)}{2!}\cdot(x-a)^2+\ldots+\frac{f^n(a)}{n!}\cdot(x-a)^n\right]+\alpha(x)\cdot(x-a)^n \text{ [V]}.$$

Really, if $x\in D_R(a)$ then the representation **[V]** is exactly the representation **[M]**. And if $x=a$, then both sides of **[V]** are equal to $f(a)$, so **[V]** is true $\forall x\in O_R(a)$. And from now on we will always use the representation **[V]** (instead of **[M]**). In general, the **Taylor's theorem** is a very powerful tool for many purposes, we will show soon the applications.

It's easy to get the Taylor's representations **[V]** of elementary functions (i.e., trigonometric, power, logarithmic functions), and it's convenient to remember most of them.

Comment. In each case here the last term $o(x^k)$ is $\alpha(x)\cdot x^k$, where $\alpha(x)$ is infinitely small when $x\to 0$ and continuous at $x=0$, i.e., $\lim_{x\to 0}\alpha(x)=0=\alpha(0)$.

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\ldots.+o(x^4)\ \forall x\in R \ \textbf{ and } \ \ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+o(x^4)\ \forall x\in(-1,1],$$

$$\sin x=x-\frac{x^3}{3!}+o(x^4)\ \forall x\in R \ \textbf{ and } \ \arcsin x=x+\frac{x^3}{6}+o(x^4)\ \forall x\in[-1,1],$$

$$tgx=x+\frac{x^3}{3}+o(x^4)\ \forall x\in(-\pi/2,\pi/2) \ \textbf{ and } \ arctgx=x-\frac{x^3}{3}+o(x^4)\ \forall x\in[-1,1],$$

$$\cos x=1-\frac{x^2}{2!}+o(x^3)\ \forall x\in R \ \textbf{ and } \ (1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+o(x^2)\,||\,\forall\alpha\in R,\forall x\in(-1,1).$$

These representations allow us to determine values of really "difficult" limits. Let's give some theory and examples.

In practice we often deal with limits like $\lim_{x\to a}f(x)^{g(x)}$, where $\lim_{x\to a}f(x)=1$ and $\lim_{x\to a}g(x)=\infty$, in such case we say that "we have a limit 1^∞" as $\lim_{x\to a}f(x)=1$, then $f(x)$ is obviously positive on some deleted neighborhood $D_\delta(a)$, and because of that $\forall x\in D_\delta(a)$ we can rewrite $f(x)^{g(x)}$ as $e^{\ln f(x)^{g(x)}}=e^{g(x)\cdot\ln f(x)}=e^{\frac{\ln f(x)}{1/g(x)}}$. Here we have the fraction $\frac{\ln f(x)}{1/g(x)}$, where numerator and denominator both go to zero when $x\to a$. Such limit $\lim_{x\to a}\frac{\ln f(x)}{1/g(x)}$ can be

calculated if we use Taylor's representations for it's functions. And after that, if $\lim_{x\to a}\dfrac{\ln f(x)}{1/g(x)}=\mathrm{A}$,

then the initial limit $\lim_{x\to a} f(x)^{g(x)}$ is equal to e^{A} (it follows from continuity of e^x at every point $\mathrm{A}\in R$).

Def. The expression $e^{v(x)}$, where $v(x)$ is some function, can be rewritten like $e^{v(x)}\equiv\exp\{v(x)\}$. It is convenient when $v(x)$ is some voluminous expression.

Example1. Calculate the limit $\lim_{x\to 0}\left(1+\sin^2 x+arctg^2 x\right)^{1/\sin x^2}$. **Answer**: e^2.

Solution. Here we have a limit 1^∞, let's rewrite it like $\exp\left\{\dfrac{\ln(1+\sin^2 x+arctg^2 x)}{\sin x^2}\right\}$ **[Ex]**.

Let's use $\sin x = x-\dfrac{x^3}{3!}+\alpha(x)\cdot x^4$, $arctgx = x-\dfrac{x^3}{3}+\beta(x)\cdot x^4$. Here and everywhere later, every function (the last term) like $\alpha(x)$ or $\beta(x)$ is such that:
$\lim_{x\to 0}\alpha(x)=0=\alpha(0)$, $\lim_{x\to 0}\beta(x)=0=\beta(0)$. Then

$$1+\sin^2 x+arctg^2 x=1+\left(x-\frac{x^3}{3!}+\alpha(x)\cdot x^4\right)^2+\left(x-\frac{x^3}{3}+\beta(x)\cdot x^4\right)^2=$$

$=1+\left(x^2-2\cdot\dfrac{x^4}{3!}+v(x)\cdot x^4\right)+\left(x^2-2\cdot\dfrac{x^4}{3}+u(x)\cdot x^4\right)$, where $v(x)$ and $u(x)$ are infinitely small when $x\to 0$ and both continuous at $x=0$. Then $1+\sin^2 x+arctg^2 x=1+2x^2-x^4+t(x)\cdot x^4$, where $t(x)$ has the same property as $v(x)$ or $u(x)$ or $\alpha(x)$ or $\beta(x)$, each new auxiliary function is a function: infinitely small when $x\to 0$ and continuous at $x=0$.

So, the numerator in **[Ex]**: $\ln(1+\sin^2 x+arctg^2 x)=\ln(1+[2x^2-x^4+t(x)\cdot x^4])$. Let's use the formula $\ln(1+x)=x+h(x)\cdot x$, then:
$\ln(1+(2x^2-x^4+t(x)\cdot x^4))=(2x^2-x^4+t(x)\cdot x^4)+h(2x^2-x^4+t(x)\cdot x^4)\cdot(2x^2-x^4+t(x)\cdot x^4)$.

The function $h(2x^2-x^4+t(x)\cdot x^4)$ is a composite function, here $2x^2-x^4+t(x)\cdot x^4$ is infinitely small when $x\to 0$ and continuous at 0, and the same is true for h, then the composite function $h(2x^2-x^4+t(x)\cdot x^4)\equiv\delta(x)$ is infinitely small when $x\to 0$ and continuous at 0,

so we can rewrite: $\ln(1+(2x^2-x^4+t(x)\cdot x^4))=(2x^2-x^4+t(x)\cdot x^4)+\delta(x)\cdot(2x^2-x^4+t(x)\cdot x^4)=$
$=(2x^2-x^4+t(x)\cdot x^4)+\delta(x)\cdot(2x^2-x^4+t(x)\cdot x^4)=2x^2+\omega(x)\cdot x^2$.

And the denominator in **[Ex]**: $\sin x^2$, let's use the formula $\sin x = x+s(x)\cdot x$, then
$\sin x^2 = x^2+s(x^2)\cdot x^4$. Similarly, the composite function $s(x^2)$ can be rewritten as $q(x)$.

And $\sin x^2 = x^2 + q(x)\cdot x^4$. From **[Ex]** we see that we are interested in the next limit:

$$\lim_{x\to 0}\frac{\ln(1+\sin^2 x + arctg^2 x)}{\sin x^2} = \lim_{x\to 0}\frac{2x^2+\omega(x)\cdot x^2}{x^2+q(x)\cdot x^4} = \lim_{x\to 0}\frac{2+\omega(x)}{1+q(x)\cdot x^2} = \frac{2}{1} = 2,$$

then the initial limit is equal to e^2.

Example2. $\lim_{x\to 0}\frac{4-3\sqrt[5]{x^3+1}-(\cos x)^{arctgx}}{x^8+\sin x^3}$. **Answer**: $-1/10$.

Solution. Here both numerator and denominator go to zero, so we have a limit $\frac{0}{0}$. Let's use Taylor's representations. The most puzzling function here is $(\cos x)^{arctgx}$, it is obviously a generalized power function, so we need to rewrite it as $(\cos x)^{arctgx} = e^{\ln(\cos x)^{arctgx}} = e^{arctgx\ln(\cos x)}$ **[S]**.

Here $\cos x = 1-\frac{x^2}{2!}+\alpha(x)\cdot x^3$, then $\ln(\cos x) = \ln\left(1-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right) = \ln\left(1+\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)\right)$

let's use the formula $\ln(1+x) = x+\beta(x)\cdot x$, then

$$\ln\left(1+\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)\right) = \left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)+\beta\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)\cdot\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right).$$

The composite function $\beta\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)$ is infinitely small when $x\to 0$ and continuous at $x=0$, then it can be rewritten as $\lambda(x)$ and the last expression can be rewritten as

$\ln\left(1+\left(-\frac{x^2}{2!}+\alpha(x)\cdot x^3\right)\right) = -\frac{x^2}{2!}+\delta(x)\cdot x^2$. Next: $arctgx = x+\omega(x)\cdot x^2$, then the power in **[S]** is

$$arctgx\cdot\ln(\cos x) = (x+\omega(x)\cdot x^2)\cdot\left(-\frac{x^2}{2!}+\delta(x)\cdot x^2\right) = -\frac{x^3}{2!}+v(x)\cdot x^3.$$

Then **[S]** is $e^{arctgx\ln(\cos x)} = e^{-\frac{x^3}{2!}+v(x)\cdot x^3}$. Let's use the formula $e^x = 1+x+o(x) \Leftrightarrow e^x = 1+x+h(x)\cdot x$.

So $e^{-\frac{x^3}{2!}+v(x)\cdot x^3} = 1+\left(-\frac{x^3}{2!}+v(x)\cdot x^3\right)+h\left(-\frac{x^3}{2!}+v(x)\cdot x^3\right)\cdot\left(-\frac{x^3}{2!}+v(x)\cdot x^3\right) = 1-\frac{x^3}{2!}+t(x)\cdot x^3.$

There is also the expression $3\sqrt[5]{x^3+1}$ (in numerator). Let's use the formula

$(1+x)^\alpha = 1+\alpha x+r(x)\cdot x$, here $\alpha = 1/5$, then $\sqrt[5]{x^3+1} = (1+x^3)^{1/5} = 1+\frac{x^3}{5}+r(x^3)\cdot x^3 =$

$1+\frac{x^3}{5}+\omega(x)\cdot x^3.$

Then the numerator of the initial limit is

$$4-3\sqrt[5]{x^3+1}-(\cos x)^{arctgx}=4-3\cdot\left(1+\frac{x^3}{5}+r(x^3)\cdot x^3\right)-\left(1-\frac{x^3}{2!}+t(x)\cdot x^3\right)=-\frac{3x^3}{5}+\frac{x^3}{2}+\lambda(x)\cdot x^3=$$

$=-\dfrac{x^3}{10}+\lambda(x)\cdot x^3$. Let's consider the denominator $x^8+\sin x^3$ as

$\sin x=x+\gamma(x)\cdot x\Rightarrow\sin x^3=x^3+\gamma(x^3)\cdot x^3$, we can rewrite it as $\sin x^3=x^3+p(x)\cdot x^3$.

So, the needed limit can be rewritten:

$$\lim_{x\to0}\frac{4-3\sqrt[5]{x^3+1}-(\cos x)^{arctgx}}{x^8+\sin x^3}=\lim_{x\to0}\frac{-\frac{x^3}{10}+\lambda(x)\cdot x^3}{x^8+(x^3+p(x)\cdot x^3)}=\lim_{x\to0}\frac{-\frac{1}{10}+\lambda(x)}{x^5+(1+p(x))}=\frac{-\frac{1}{10}}{1}=-\frac{1}{10}$$

Taylor's representations also allow us to find approximate values of functions.

Example3. Let's calculate $\cos10^o$ with the accuracy $\varepsilon=0{,}002$.

Solution. We need to convert degrees in radians $\cos10^o=\cos\dfrac{\pi}{18}$. Let's use the formula **[T]** for $\cos x$, so there exist the point c in the interval with ends $0,x$ such that:

$$\cos x=1-\frac{x^2}{2!}+\frac{\cos_{xxx}(c)}{3!}x^3\Leftrightarrow\cos x=1-\frac{x^2}{2!}+\frac{\sin(c)}{3!}x^3$$

(we are free to choose how many terms we want to take in **[T]**).

From here we have: $\cos x-\left(1-\dfrac{x^2}{2!}\right)=\dfrac{\sin(c)}{3!}x^3\Rightarrow\left|\cos x-\left(1-\dfrac{x^2}{2!}\right)\right|=\left|\dfrac{\sin(c)}{3!}x^3\right|\le\dfrac{x^3}{6}$,

if $x=\dfrac{\pi}{18}<\dfrac{3{,}5}{18}<0{,}2$, then $\dfrac{x^3}{6}<\dfrac{(0{,}2)^3}{6}<0{,}002$. Then for $x=\dfrac{\pi}{18}$ we have $\left|\cos x-\left(1-\dfrac{x^2}{2!}\right)\right|<0{,}002$.

So when $x=\dfrac{\pi}{18}$, the expression $\left(1-\dfrac{x^2}{2!}\right)$ approximates $\cos x$ with $\varepsilon=0{,}002$ accuracy.

Then $\cos\dfrac{\pi}{18}\approx1-\dfrac{(\pi/18)^2}{2!}\Leftrightarrow\cos\dfrac{\pi}{18}\approx1-\dfrac{\pi^2}{648}$.

Taylor's formulas **[T]** and **[M]** has a great importance in mathematics. When we use these formulas, especially for limit-calculation, we should be careful and do not lose any terms. In some cases we can use more simple tool in order to calculate some "difficult" limits, it is called a **L'Hospital rule**.
Notice that the **L'Hospital rule** is a great tool, but it doesn't work in all cases. If we apply this rule to the previous **examples1,2** we will not get any result. Taylor's representation is a much stronger tool, but the use of it requires from us to perform much more work.

L'Hospital rule. We want to calculate some limit $\lim_{x\to a}\dfrac{f(x)}{g(x)}$, where $f(x)$ and $g(x)$ are defined and differentiable on some deleted neighborhood $D_R(a)$ and $g(x)\neq 0,\ g_x(x)\neq 0$ on $D_R(a)$. And we have the situation: $\lim_{x\to a} f(x)=0$ and $\lim_{x\to a} g(x)=0$. So the limit $\lim_{x\to a}\dfrac{f(x)}{g(x)}$ is a limit $\dfrac{0}{0}$. In such case we can consider the limit $\lim_{x\to a}\dfrac{f'(x)}{g'(x)}$, if this limit exists, then the initial limit $\lim_{x\to a}\dfrac{f(x)}{g(x)}$ also exists and these limits are equal.

In practice this tool allows us to solve problems very quickly.

Example. Find $\lim_{x\to 0}\dfrac{\sin mx}{\sin nx}$. Here we see a limit $\dfrac{0}{0}$, let's calculate $\lim_{x\to 0}\dfrac{(\sin mx)'}{(\sin nx)'}=\lim_{x\to 0}\dfrac{\cos mx\cdot m}{\cos nx\cdot n}=\dfrac{m}{n}$, then the initial limit $\lim_{x\to 0}\dfrac{\sin mx}{\sin nx}$ is $\dfrac{m}{n}$.

Let's prove the **L'Hospital rule**. We define two auxiliary functions $\tilde{f}(x)$ and $\tilde{g}(x)$:
$\tilde{f}(x)\equiv f(x)\ \forall x\in D_R(a)$ and $\tilde{f}(a)\equiv //\, by\ def\, //\equiv 0$ and similarly:
$\tilde{g}(x)\equiv g(x)\ \forall x\in D_R(a)$ and $\tilde{g}(a)\equiv //\, by\ def\, //\equiv 0$.

The functions $\dfrac{f(x)}{g(x)}$ and $\dfrac{\tilde{f}(x)}{\tilde{g}(x)}$ are both defined and equal on $D_R(a)$, then the existence of $\lim_{x\to a}\dfrac{\tilde{f}(x)}{\tilde{g}(x)}$ is tantamount to the existence of $\lim_{x\to a}\dfrac{f(x)}{g(x)}$. If one of these limit's exist, then the other also exists and these limits are equal. Let's assume that $\lim_{x\to a}\dfrac{f'(x)}{g'(x)}=\mathrm{A}$.

Let's fix an arbitrary sequence $\{x_n\}\subset D_R(a)\,||\,\{x_n\}\to a$ if we show that $\left\{\dfrac{\tilde{f}(x_n)}{\tilde{g}(x_n)}\right\}\to \mathrm{A}$, then $\lim_{x\to a}\dfrac{\tilde{f}(x)}{\tilde{g}(x)}=\mathrm{A}$, and it is exactly what we need.

Both functions $\tilde{f}(x)$ and $\tilde{g}(x)$ are differentiable on $D_R(a)$. For every x_n we have:
$\dfrac{\tilde{f}(x_n)}{\tilde{g}(x_n)}=\dfrac{\tilde{f}(x_n)-\tilde{f}(a)}{\tilde{g}(x_n)-\tilde{g}(a)}=\begin{bmatrix} Cauchy \\ formula\end{bmatrix}=\dfrac{\tilde{f}'(c_n)}{\tilde{g}'(c_n)}$. As $\{x_n\}\to a$, there must be $\{c_n\}\to a$ (because every c_n is between x_n and a). As we have $\lim_{x\to a}\dfrac{f'(x)}{g'(x)}=\mathrm{A}$, then $\left\{\dfrac{f'(c_n)}{g'(c_n)}\right\}\to \mathrm{A}$, which is equivalent

to $\left\{\dfrac{\tilde{f}(x_n)}{\tilde{g}(x_n)}\right\} \to A$, everything is proved.

Comment. L'Hospital rule can be applied repeatedly, i.e., we can apply it for $\lim_{x\to a}\dfrac{f'(x)}{g'(x)}$ if necessary (and if possible) and then for $\lim_{x\to a}\dfrac{f''(x)}{g''(x)}$ if necessary and etc.

For Reader's practice:

[1] Find the limit $\lim_{n\to\infty} n\cos\left(\dfrac{\pi}{4n}\right)\sin\left(\dfrac{\pi}{4n}\right)$. **Answer:** $\dfrac{\pi}{4}$.

[2] Find the limit $\lim_{x\to 0}\dfrac{a^x - b^x}{x}$. **Answer:** $\ln\dfrac{a}{b}$.

[3] Find the limit $\lim_{x\to 1}\dfrac{\left(\sum_{k=1}^{100} x^k\right) - 100}{x-1}$. **Answer:** 5050.

[4] Find the limit $\lim_{x\to 0}\left(\dfrac{1}{x^2} - ctg^2 x\right)$ **Answer:** $\dfrac{2}{3}$.

[5] Find the limit $\lim_{x\to\infty}\left(tg\,\dfrac{\pi x}{2x+1}\right)^{1/x}$ **Answer:** 1.

[6] Find the limit $\lim_{x\to 0}\left(\dfrac{1}{\sin x\cdot arctgx} - \dfrac{1}{tgx\cdot\arcsin x}\right)$ **Answer:** 1.

[7] Find the limit $\lim_{x\to 0}\left(\dfrac{\sqrt{1+x}\cdot\sin x + \ln(\cos x) - x}{\sqrt[3]{1-x^3} - 1}\right)$ **Answer:** $\dfrac{7}{8}$.

[8] Find the limit $\lim_{x\to 0}\left(\dfrac{\sqrt{1+2x^3} - \cos x^4}{tgx - x}\right)$ **Answer:** 3.

[8] Find the limit $\lim_{x\to 0}\left(\dfrac{\sqrt{1+2tgx} - e^x + x^2}{\arcsin x - \sin x}\right)$ **Answer:** 2

[9] Find the limit $\lim_{x\to 0}\left(\dfrac{xe^{tgx} - \sin^2 x - x}{x + x^3 - tgx}\right)$ **Answer:** $\dfrac{3}{4}$.

Exercise*: $\{x_n\} \to a$, show that **[1]** the sequence $\left\{\dfrac{x_1 + x_2 + \ldots + x_n}{n}\right\}$ also goes to a.

[2] If $\{x_n\}$ is a positive sequence, then $\{\sqrt[n]{x_1\cdot x_2\cdot\ldots\cdot x_n}\}$ goes to a.

Integral

From now on, $f(x)$ is any function that is defined on $[a,b]$ and **bounded** on $[a,b]$.

Let's divide the segment $[a,b]$ into several consecutive segments $\Delta_1, \Delta_2 .... \Delta_n$, we choose an arbitrary point ξ_k on every segment Δ_k, the sum $\Sigma \equiv f(\xi_1)\cdot\Delta_1 + f(\xi_2)\cdot\Delta_2 + ... + f(\xi_n)\cdot\Delta_n$ is called an integral sum of f with respect to the partition $\Delta_1, \Delta_2 ... \Delta_n \equiv \{\Delta_k\}$. This sum depends on the points $\xi_k \in \Delta_k \,||\, k = 1..\, n$. The maximal length $\max\{\Delta_k\}$ (i.e., the length of the longest segment) is called a norm of the partition $\{\Delta_k\}$.

Def. The number A is called an integral f on $[a,b]$ if for any (small) positive $\varepsilon > 0$ we can find the positive $\delta > 0$, such that for any partition $\{\Delta_k\}$ of $[a,b]$ such that $\max\{\Delta_k\} < \delta$, for any points $\xi_k \in \Delta_k \,||\, k = 1..\, n$, the integral sum $\Sigma \equiv f(\xi_1)\cdot\Delta_1 + f(\xi_2)\cdot\Delta_2 + ... + f(\xi_n)\cdot\Delta_n$ is "ε close to A", i.e., $|\mathrm{A} - \Sigma| < \varepsilon$.

In this case we denote $\mathrm{A} = \int_a^b f(x)dx$ and we say that f is integrable on $[a,b]$, and A is an integral of f on $[a,b]$.

Exercise1. Show that if such number $\mathrm{A} = \int_a^b f(x)dx$ exists, than it is unique (there can't be two different numbers $\mathrm{A} \neq \mathrm{B}$ which are both integrals of f on $[a,b]$).

As f is bounded on $[a,b]$, it is bounded on every segment Δ_k, so the set of all values $\{f(x) \,||\, x \in \Delta_k\}$ has a supremum M_k and an infinum m_k. The sum $\Sigma^{low} \equiv m_1\cdot\Delta_1 + m_2\cdot\Delta_2 + ... + m_n\cdot\Delta_n$ is called a lower (integral) sum of f with respect to $\{\Delta_k\}$ and the sum $\Sigma^{up} \equiv \mathrm{M}_1\cdot\Delta_1 + \mathrm{M}_2\cdot\Delta_2 + ... + \mathrm{M}_n\cdot\Delta_n$ is called an upper (integral) sum of f with respect to $\{\Delta_k\}$.

Notice. For any concrete partition $\{\Delta_k\}$, a lower and an upper integral sums Σ^{low} and Σ^{up} are concrete real numbers, these numbers are uniquely defined by the partition $\{\Delta_k\}$. And an ordinary integral sum $\Sigma \equiv f(\xi_1)\cdot\Delta_1 + f(\xi_2)\cdot\Delta_2 + ... + f(\xi_n)\cdot\Delta_n$ is not uniquely defined by $\{\Delta_k\}$, this sum is defined by the choice of points $\xi_k \in \Delta_k \,||\, k = 1..\, n$.

For any partition $\{\Delta_k\}$ we obviously have $\Sigma^{low} \leq \Sigma \leq \Sigma^{up}$.

Exersice2. Let's fix any partition $\{\Delta_k\}$.

We have the set $\{\Sigma\}$ of all integral sums with respect to $\{\Delta_k\}$. Show that the upper sum Σ^{up} is a supremum of $\{\Sigma\}$, and the lower sum Σ^{low} is an infinum of $\{\Sigma\}$.

Theorem1 [Integrability criterion]. f is integrable on $[a,b] \Leftrightarrow$ for any $\varepsilon > 0$ there exist $\delta > 0$ such that for any partition $\{\Delta_k\}$, which norm is less than δ, we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$.

Proof. $\Rightarrow$Let f is integrable on $[a,b]$. We fix an arbitrary positive $\varepsilon > 0$, for the positive number $\varepsilon / 4 > 0$ there exist some $\delta > 0$ such that for any partition $\{\Delta_k\}$, which norm is less than δ, for any integral sum $\Sigma \equiv f(\xi_1) \cdot \Delta_1 + f(\xi_2) \cdot \Delta_2 + \ldots + f(\xi_n) \cdot \Delta_n$ we have $|A - \Sigma| < \varepsilon / 4$.
Let's fix any such partition. Any integral sum Σ lies inside the interval $(A - \varepsilon/4, A + \varepsilon/4)$.
Then the set of all integral sums $\{\Sigma\}$ lies inside $(A - \varepsilon/4, A + \varepsilon/4)$. As Σ^{up} is a supremum of the set $\{\Sigma\} \subset (A - \varepsilon/4, A + \varepsilon/4)$, then there must be $\Sigma^{up} \in [A - \varepsilon/4, A + \varepsilon/4]$, as Σ^{low} is an infinum of the set $\{\Sigma\} \subset (A - \varepsilon/4, A + \varepsilon/4)$, then there must be $\Sigma^{low} \in [A - \varepsilon/4, A + \varepsilon/4]$. So Σ^{low} and Σ^{up} are some numbers from the segment $[A - \varepsilon/4, A + \varepsilon/4]$, then the difference $\Sigma^{up} - \Sigma^{low}$ is not greater than the length of this segment, so $\Sigma^{up} - \Sigma^{low} < \varepsilon/2 < \varepsilon$.

We had started from an arbitrary positive $\varepsilon > 0$, and we found $\delta > 0$ such that for any partition, which norm is less than δ, we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$.

Conversely. $\Leftarrow$ Let's fix an arbitrary positive $\varepsilon > 0$, there exist $\delta > 0$ such that for any partition $\{\Delta_k\}$, which norm is less than δ, we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$.

Auxiliary1. Let's notice an important fact. Any lower integral sum $\tilde{\Sigma}^{low}$ is not greater than any upper integral sum Σ^{up}, i.e., $\tilde{\Sigma}^{low} \le \Sigma^{up}$ even if these sums correspond to different partitions $\{\tilde{\Delta}_k\}, \{\Delta_k\}$ of $[a,b]$. Really, let's fix an arbitrary partition $\{\Delta_k\}$, suppose that we added only one new point to $\{\Delta_k\}$, so it divides some segment Δ_j in two segments $\tilde{\Delta}_j, \tilde{\Delta}_{j+1}$.

Let's consider f on Δ_j. We obviously have $M_j \equiv \sup_{x \in \Delta_j} f(x) \ge \tilde{M}_j \equiv \sup_{x \in \tilde{\Delta}_j} f(x)$, because $\Delta_j \supset \tilde{\Delta}_j$. And similarly $M_j \equiv \sup_{x \in \Delta_j} f(x) \ge \tilde{M}_{j+1} \equiv \sup_{x \in \tilde{\Delta}_{j+1}} f(x)$ (because $\Delta_j \supset \tilde{\Delta}_{j+1}$).

From here follows that

$$\left(\tilde{M}_j \cdot \tilde{\Delta}_j + \tilde{M}_{j+1} \cdot \tilde{\Delta}_{j+1}\right) \le \left(M_j \cdot \tilde{\Delta}_j + M_j \cdot \tilde{\Delta}_{j+1}\right) = M_j \cdot (\tilde{\Delta}_j + \tilde{\Delta}_{j+1}) = M_j \cdot \Delta_j.$$

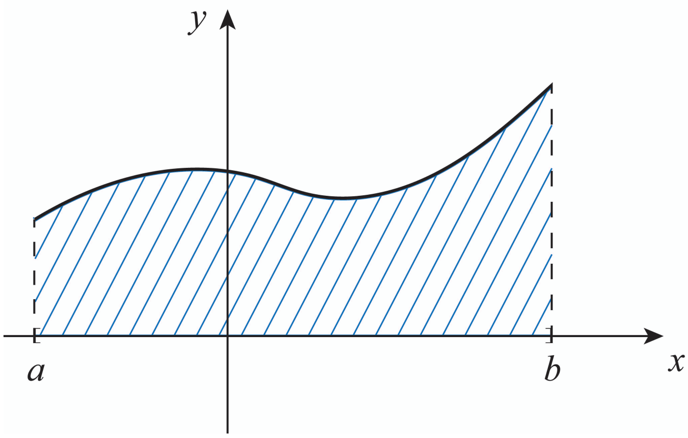
From here follows that the upper sum $\tilde{\Sigma}^{up}$ of f with respect to the new partition (where one new point is added) is not greater than the initial upper sum Σ^{up} of f with respect to the old partition.

Conclusion. When we add some new points to our partition $\{\Delta_k\}$ (i.e., we make a "refinement of it") an upper integral sum Σ^{up} may only become less. And it's very easy to understand (similarly, by adding one new point to the initial partition) that a lower integral sum Σ^{low} may only become greater (after any refinement).

Let now we have two different upper and lower integral sums $\widetilde{\Sigma}^{low}, \Sigma^{up}$ which are built with respect to some partitions $\{\widetilde{\Delta}_k\},\{\Delta_k\}$. Let's unite these partitions $\{\widetilde{\Delta}_k\},\{\Delta_k\}$ (we just add all the points of one partition to other partition). The new partition $\{\overline{\Delta}_k\}$ is a refinement of $\{\widetilde{\Delta}_k\}$ and it is also a refinement of $\{\Delta_k\}$, let $\overline{\Sigma}^{up}, \overline{\Sigma}^{low}$ are the upper and lower integral sums with respect to $\{\overline{\Delta}_k\}$. Then we have $\widetilde{\Sigma}^{low} \le \overline{\Sigma}^{low}$ and $\overline{\Sigma}^{up} \le \Sigma^{up}$ and obviously $\overline{\Sigma}^{low} \le \overline{\Sigma}^{up}$, then $\widetilde{\Sigma}^{low} \le \Sigma^{up}$.
And the auxiliary1 is proved.

Let's finish the proof of the converse assertion (theorem1). We consider the set of all possible lower sums $\{\Sigma^{low}\}$ (every sum is built with respect to some partition of $[a,b]$) and the set of all possible upper sums $\{\Sigma^{up}\}$. Any element of $\{\Sigma^{low}\}$ is not greater than any element of $\{\Sigma^{up}\}$, then $\{\Sigma^{low}\}$ is bounded above and there exist $\sup\{\Sigma^{low}\}$, and $\{\Sigma^{up}\}$ is bounded below and there exist $\inf\{\Sigma^{up}\}$. From the initial condition, for any small $\varepsilon > 0$ there exist some concrete elements $\Sigma^{low} \in \{\Sigma^{low}\}$ and $\Sigma^{up} \in \{\Sigma^{up}\}$ such that $\Sigma^{up} - \Sigma^{low} < \varepsilon$. From here follows that $\sup\{\Sigma^{low}\} \equiv \inf\{\Sigma^{up}\} \equiv \mathrm{A}$.
Let's show that A is an integral of f on $[a,b]$. We fix an arbitrary positive $\varepsilon > 0$, there exist $\delta > 0$ such that for any partition $\{\Delta_k\}$, which norm is less than δ, we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$. From the definition of A follows that $\Sigma^{low} \le \mathrm{A} \le \Sigma^{up}$, and for any integral sum Σ, which is built with respect to $\{\Delta_k\}$, we also have $\Sigma^{low} \le \Sigma \le \Sigma^{up}$. Then both numbers A, Σ belong to the segment $[\Sigma^{low}, \Sigma^{up}]$, the length of this segment is less than ε, then $|\mathrm{A} - \Sigma| < \varepsilon$.Then, by definition, $\mathrm{A} = \int\limits_a^b f(x)dx$.

Geometrical meaning. For any positive function f on $[a,b]$, if f is integrable on $[a,b]$, then $\int\limits_a^b f(x)dx$ is an area of the figure Ω which "lies below" the graph of f **[pict1]**.



pict.1

Let's fix an arbitrary partition of $[a,b]$.

Any upper sum Σ^{up} is an area of a figure Ω^{ext}, which consists of rectangles and contains Ω.
And any lower sum Σ^{low} is an area of a figure Ω^{int}, which consists of rectangles and belongs to Ω. So we have $\Omega^{\text{int}} \subset \Omega \subset \Omega^{ext}$ and $\Sigma^{low} = S(\Omega^{\text{int}})$ and $\Sigma^{up} = S(\Omega^{ext})$. According to the theorem1[**Integrability criterion**], for any $\varepsilon > 0$ we can fix some measurable figures $\Omega^{\text{int}} \subset \Omega \subset \Omega^{ext}$ such that $S(\Omega^{ext}) - S(\Omega^{\text{int}}) < \varepsilon$.

From here (assertion4 **[1-st criterion of measurability]** "Area construction", Book 1) follows that Ω is measurable, and the number $S(\Omega)$ is defind. Let's fix now any positive sequence $\{\varepsilon_n\} \to 0$, for any number ε_n let's fix the internal figure $\Omega^{\text{int}}(n)$ and the external figure $\Omega^{ext}(n)$ such that $\left|S(\Omega^{ext}(n)) - S(\Omega^{\text{int}}(n))\right| < \varepsilon_n$ **[T]**. Then we have the sequence of internal figures $\{\Omega^{\text{int}}(n)\} \,||\, \Omega^{\text{int}}(n) \subset \Omega$ and the sequence of external figures $\{\Omega^{ext}(n)\} \,||\, \Omega^{ext}(n) \supset \Omega$.

As $\Omega^{\text{int}}(n) \subset \Omega \subset \Omega^{ext}(n)$, then $S(\Omega^{\text{int}}(n)) \le S(\Omega) \le S(\Omega^{ext}(n)) \,||\, \forall n$ **[E]**.

In the integrability criterion (theorem1) we showed that $\mathrm{A} = \int_a^b f(x)dx$ is a supremum of the set $\{\Sigma^{low}\}$ of all lower integral sums and an infinum of the set $\{\Sigma^{up}\}$ of all upper integral sums.

So, **for any possible** lower and upper sums we have $\Sigma^{low} \le \mathrm{A} \le \Sigma^{up}$. **Notice** that every area $S(\Omega^{\text{int}}(n))$ is some lower integral sum (by construction), and every area $S(\Omega^{ext}(n))$ is some upper integral sum Σ^{up}. Then there must be $S(\Omega^{\text{int}}(n)) \le \mathrm{A} \le S(\Omega^{ext}(n)) \,||\, \forall n$ **[E1]**. From **[E]** and **[E1]** follows that both numbers A and $S(\Omega)$ belong to the segment $[S(\Omega^{\text{int}}(n)),\ S(\Omega^{ext}(n))] \,||\, \forall n$, from **[T]** follows that the length of that segment can be less than any positive number, then

$$\mathrm{A} = S(\Omega) \Leftrightarrow S(\Omega) = \int_a^b f(x)dx.$$

Theorem2. If f is continuous on $[a,b]$, then f is integrable on $[a,b]$.

Proof. As f is continuous on $[a,b]$, then it is bounded on $[a,b]$ and we can raise a question about it's integrability. According to the Cantor's theorem, f is uniformly continuous on $[a,b]$.

Let's fix an arbitrary positive $\varepsilon > 0$. We take the positive number $\varepsilon/(b-a) > 0$, there exist $\delta > 0$ such that for any $x_1, x_2 \in [a,b]$, if $|x_1 - x_2| < \delta$, then $|f(x_1) - f(x_2)| < \varepsilon/(2\cdot(b-a))$.

Let's fix any partition $\{\Delta_n\}$ of $[a,b]$ which norm is less than δ. Let's consider f on any segment Δ_k, for any $x_1, x_2 \in \Delta_k$ we obviously have $|f(x_1) - f(x_2)| < \varepsilon/(2\cdot(b-a))$, from here immediately follows that $\mathrm{M}_k - m_k \le \varepsilon/(2\cdot(b-a))$ where $\mathrm{M}_k = \sup_{\Delta_k} f(x)$ and $m_k = \inf_{\Delta_k} f(x)$.

Let's estimate the difference between upper and lower integral sums:

$$\Sigma^{up} - \Sigma^{low} = (\mathrm{M}_1 - m_1)\cdot\Delta_1 + (\mathrm{M}_2 - m_2)\cdot\Delta_2 + \ldots + (\mathrm{M}_n - m_n)\cdot\Delta_n \le \frac{\varepsilon}{2\cdot(b-a)}\cdot\Delta_1 + \ldots + \frac{\varepsilon}{2\cdot(b-a)}\cdot\Delta_n =$$

$$= \frac{\varepsilon}{2\cdot(b-a)}\cdot(\Delta_1 + \ldots + \Delta_n) = \frac{\varepsilon}{2\cdot(b-a)}\cdot(b-a) = \frac{\varepsilon}{2} < \varepsilon.$$ So we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$.

Let's sum up, for an arbitrary positive $\varepsilon > 0$ there exist $\delta > 0$ such that for any partition $\{\Delta_n\}$, which norm is less than δ, we have $\Sigma^{up} - \Sigma^{low} < \varepsilon$, then according to the theorem1[**Integrability criterion**], f is integrable on $[a,b]$.

Theorem3. f is monotonic on $[a,b]$, then f is integrable on $[a,b]$.

Proof. Let f is monotonically increasing on $[a,b]$ (i.e., $\forall x_1 < x_2 \in [a,b] \Rightarrow f(x_1) \le f(x_2)$). Then for any $x \in [a,b]$ we have $f(a) \le f(x) \le f(b)$, then f is bounded on $[a,b]$ and we can raise a question about it's integrability on $[a,b]$. If $f(a) = f(b)$, then $f \equiv const$ on $[a,b]$ and such function is obviously integrable on $[a,b]$. Let $f(a) < f(b)$, then we fix an arbitrary positive $\varepsilon > 0$, and we take the number $\delta \equiv \varepsilon / (f(b) - f(a))$. Let's fix any partition $\{\Delta_n\}$ of $[a,b]$ which norm is less than δ.

Any value $\mathrm{M}_k \equiv \sup_{\Delta_k} f(x)$ is obviously the value of f at the right end of Δ_k.

Any value $m_k \equiv \inf_{\Delta_k} f(x)$ is obviously the value of f at the left end of Δ_k.

So let's denote $\Delta_1 \equiv [a, x_1], \Delta_2 \equiv [x_1, x_2], \Delta_3 \equiv [x_2, x_3] \ldots.. \Delta_n \equiv [x_n, b]$, then

$$\Sigma^{up} - \Sigma^{low} = (f(x_1) - f(a)) \cdot \Delta_1 + (f(x_2) - f(x_1)) \cdot \Delta_2 + \ldots. + (f(b) - f(x_n)) \cdot \Delta_n <$$

$$< (f(x_1) - f(a)) \cdot \frac{\varepsilon}{f(b) - f(a)} + (f(x_2) - f(x_1)) \cdot \frac{\varepsilon}{f(b) - f(a)} + \ldots. + (f(b) - f(x_{n-1})) \cdot \frac{\varepsilon}{f(b) - f(a)} =$$

$$= \frac{\varepsilon}{f(b) - f(a)} \cdot \left([f(x_1) - f(a)] + [f(x_2) - f(x_1)] + [f(x_3) - f(x_2)] + \ldots. + [f(b) - f(x_{n-1})] \right) =$$

$= \dfrac{\varepsilon}{f(b) - f(a)} \cdot (f(b) - f(a)) = \varepsilon$. And according to the **theorem1**[**Integrability criterion**], f is integrable on $[a,b]$.

Exercise3. By using the initial definition of an integral prove the next properties:

[1] f and g are integrable on $[a,b]$. Then $f(x) + g(x)$ is integrable on $[a,b]$ and

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

[2] f is integrable on $[a,b]$ and $\lambda \in R$. Then $\lambda \cdot f(x)$ is integrable on $[a,b]$ and

$$\int_a^b \lambda \cdot f(x)dx = \lambda \cdot \int_a^b f(x)dx.$$

From **[1]** and **[2]** follows that: f and g are integrable on $[a,b]$. Then any linear combination of these functions $\lambda \cdot f(x) + \mu \cdot g(x)$ is integrable on $[a,b]$ and

$\displaystyle\int_a^b \left(\lambda \cdot f(x) + \mu \cdot g(x)\right)dx = \lambda \cdot \int_a^b f(x)dx + \mu \cdot \int_a^b g(x)dx$, in particular $\displaystyle\int_a^b \left(f(x) - g(x)\right)dx = \int_a^b f(x)dx - \int_a^b g(x)dx$

[3] f is integrable on $[a,b]$. Then for any $c\in(a,b)$: f is integrable on $[a,c]$ and on $[c,b]$ and $\int_a^c f(x)dx+\int_c^b f(x)dx=\int_a^b f(x)dx$.

[4] f and g are integrable on $[a,b]$ and $g(x)\le f(x)$ on $[a,b]$. Then $\int_a^b f(x)dx\le\int_a^b g(x)dx$.

[5] If f is integrable on $[a,b]$. Then $|f|$ is integrable on $[a,b]$ and $\left|\int_a^b f(x)dx\right|\le\int_a^b |f(x)|dx$.

Comment. The converse assertion is not true, if $|f|$ is integrable on $[a,b]$, then f mustn't be integrable on $[a,b]$. Consider the "Dirichlet function": $f(x)\equiv 1$ when x is rational and $f(x)\equiv 0$ when x is irrational.

Mean Value theorem. f is integrable on $[a,b]$. Then there exist the number μ: $\inf_{[a,b]} f=m\le\mu\le M=\sup_{[a,b]} f$ such that $\int_a^b f(x)dx=\mu\cdot(b-a)$.

Proof. For any $x\in[a,b]$ we have

$$m\le f(x)\le M\Rightarrow\int_a^b mdx\le\int_a^b f(x)dx\le\int_a^b Mdx\Leftrightarrow m\cdot(b-a)\le\int_a^b f(x)dx\le M\cdot(b-a).$$

Then really $\int_a^b f(x)dx=\mu\cdot(b-a)$ (for some μ: $m\le\mu\le M$).

Consequence1. If f is continuous on $[a,b]$, then it reaches all it's intermediate values on $[a,b]$, then there exist $c\in[a,b]$ such that $f(c)=\mu$, and the mean value theorem looks like $\int_a^b f(x)dx=f(c)\cdot(b-a)$.

Let's extend the definition of an integral: **[1]** We define: $\int_a^a f(x)dx\equiv 0$, **[2]** f is integrable on $[a,b]$ (here $a<b$, as in all the previous cases). We define $\int_b^a f(x)dx\equiv //\,by\ def\,//\equiv-\left(\int_a^b f(x)dx\right)$.

Now: if f is integrable on some segment with ends a,b, then both integrals $\int_a^b f(x)dx$ and $\int_b^a f(x)dx$ are defined. And now the symbol $\int_a^b f(x)dx$ makes sense in all the cases: $a<b,\ a=b,\ b<a$.

It's very easy to check that the properties **[1]**,**[2]** are still true for the new "extended" integral. And **[3]** becomes true for any points a,b,c (if only all the integrals from **[3]** exist). Also, the **Mean value theorem** and the consequence1 from it are still true.

Let's sum up: the mean value theorem, and all the "linear properties" of integral are true in any case $a<b,\ a=b,\ b<a.$

But remember, that all the properties concerning the estimation of integral values (like **[4]** and **[5]**) are true only when $a \le b$ (it's very easy to understand why). So, when we estimate some integral value, we need to make sure at first that $a \le b$.

Def. $f(x)$ is defined on $[a,b]$. If there exist the function $F(x)$,which is differentiable on $[a,b]$, such that $F_x(x)=f(x)\ \forall x\in[a,b]$ (at the end points a,b we imply the left and right derivatives of F), then $F(x)$ is called an antiderivative of $f(x)$ on $[a,b]$.

The symbol $F(x)=\int f(x)dx$ means that $F(x)$ is an antiderivative of $f(x)$ on some segment.

Assertion1. If $F(x)$ is an antiderivative of $f(x)$ on $[a,b]$, then for any constant $C\in R$ the function $F(x)+C$ is also an antiderivative of $f(x)$ on $[a,b]$. And if $F(x)$ and $G(x)$ are antiderivatives of $f(x)$ on $[a,b]$, then there exist some constant C such that

$F(x)=G(x)+C \,||\, \forall x\in[a,b]$.

The first part is obvious. In the second part we need to consider the auxiliary function $\varphi(x)\equiv F(x)-G(x)$, we have $\varphi_x(x)=F_x(x)-G_x(x)\equiv 0$ on $[a,b]$, then $\varphi_x(x)=const$ on $[a,b]$.

Fundamental theorem of calculus.

$f(x)$ is continuous on $[a,b]$, then $F(x)\equiv\int\limits_a^x f(t)dt$ is an antiderivative of $f(x)$ on $[a,b]$.

Comment. For any concrete number $x\in[a,b]$ the integral $\int\limits_a^x f(t)dt$ exists (because f is continuous on $[a,x]$). So, for any $x\in[a,b]$ the value $\int\limits_a^x f(t)dt$ is a concrete real number, then $F(x)\equiv\int\limits_a^x f(t)dt$ is really a function of variable x on $[a,b]$. We need to prove that at any point $x_0\in[a,b]$ we have $F_x(x_0)=f(x_0)$.

Proof. We fix an arbitrary point $x_0\in[a,b]$ and we consider the function $F(x)\equiv\int\limits_a^x f(t)dt$, the derivative of this function at x_0 is the next limit $\lim_{x\to x_0}\dfrac{F(x)-F(x_0)}{x-x_0}$ (if this limit exists).

Let's consider $\frac{F(x)-F(x_0)}{x-x_0}=\frac{\int\limits_a^x f(t)dt-\int\limits_a^{x_0} f(t)dt}{x-x_0}$ **[V]**. It's easy to notice that

$\int\limits_a^{x_0} f(t)dt+\int\limits_{x_0}^x f(t)dt=\int\limits_a^x f(t)dt$ (it doesn't matter $x_0<x$ or $x<x_0$). Then **[V]** can be rewritten like

$\frac{\int\limits_{x_0}^x f(t)dt}{x-x_0}$ **[V1]**. Let's apply now the mean value theorem for f on the segment with the ends x_0, x,

as f is continuous, there exist some point $c=c(x)$ such that $\int\limits_{x_0}^x f(t)dt=f(c(x))\cdot(x-x_0)$.

Notice, we have denoted $c=c(x)$ as a function of x, because every $x\neq x_0$ from $[a,b]$ defines some point (number) $c=c(x)$ which lies in the interval with ends x_0 and x. Then **[V1]** can be rewritten as $\frac{f(c(x))\cdot(x-x_0)}{x-x_0}=f(c(x))$. And therefore $\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0}=\lim_{x\to x_0} f(c(x))$ **[V2]**.

Obviously $\lim_{x\to x_0} c(x)=x_0$ (because $c(x)$ is always between x and x_0), as $f(x)$

is continuous at x_0, the limit of the composite function $\lim_{x\to x_0} f(c(x))$ exists and equals $f(x_0)$ (Improvement of Theorem3 [**Limit of a composite function**]).

And from **[V2]** we have $\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0}=f(x_0)$. Everything is proved.

Consecuence2. From the assertion1 follows that any antiderivative of a continuous function $f(x)$

on $[a,b]$ must look like $F(x)=\int\limits_a^x f(t)dt+C$, where $C\equiv const$.

Consecuence3 [**Newton-Leibniz formula**]. $f(x)$ is continuous on $[a,b]$ and $F(x)$ is any

antiderivative of $f(x)$ on $[a,b]$, then $\int\limits_a^b f(t)dt=F(b)-F(a)$.

Proof. Any antiderivative looks like $F(x)=\int\limits_a^x f(t)dt+C$, where $C=const$, then

$$F(b)-F(a)=\left(\int\limits_a^b f(t)dt+C\right)-\left(\int\limits_a^a f(t)dt+C\right)=\left(\int\limits_a^b f(t)dt+C\right)-(0+C)=\int\limits_a^b f(t)dt.$$

From here follows that the [Newton-Leibniz formula] can be applied not only when $a<b$, but also when $a>b$ or $a=b$. So we can apply this formula for any kind of integral.

[Newton-Leibniz formula]: $f(x)$ is continuous on some segment with ends a,b and $F(x)$ is any antiderivative of $f(x)$ on that segment, then $\int_a^b f(t)dt = F(b)-F(a)$.

Consecuence4. f is differentiable on a segment with ends a,b, then $f(x)$ is an antiderivative of $f_x(x)$ on that segment and $\int_a^b f_x(t)dt = f(b)-f(a)$.

Def. A function is called smooth on $[a,b]$ if it's derivative (as an independent function) is continuous on $[a,b]$.

Theorem4 [Change of a variable]. $f(x)$ is continuous on $[a,b]$. And $x=\varphi(t)$ is smooth on $[\alpha,\beta]$, all the values of $\varphi(t)$ on $[\alpha,\beta]$ belong to $[a,b]$ and $\varphi(\alpha)=a,\ \varphi(\beta)=b$.

Then $\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\cdot\varphi_t(t)dt$.

Proof. From the initial condition follows that both integrals $\int_a^b f(x)dx$ and $\int_\alpha^\beta f(\varphi(t))\cdot\varphi_t(t)dt$ do exist (because these are integrals of continuous functions). Let's fix any antiderivative $F(x)$ of $f(x)$ on $[a,b]$. And let's consider the function $F(\varphi(t))$, it is a composite function, and it is defined on $[\alpha,\beta]$. For any point $t_0\in[\alpha,\beta]$, the function $x=\varphi(t)$ is differentiable at t_0, and $F(x)$ is differentiable at $x_0=x(t_0)$ (because F is differentiable everywhere on $[a,b]$, as $F_x(x)\equiv f(x)$ by definition), then the composite function $F(\varphi(t))$ is differentiable at t_0, and there must be $F_t(\varphi(t_0)) = F_x(\varphi(t_0))\cdot\varphi_t(t_0) = [as\ F_x(x)\equiv f(x)] = f(\varphi(t_0))\cdot\varphi_t(t_0)$.

So we got the formula $F_t(\varphi(t_0)) = f(\varphi(t_0))\cdot\varphi_t(t_0) \,||\, \forall t_0\in[\alpha,\beta]$, we can rewrite it as $F_t(\varphi(t)) = f(\varphi(t))\cdot\varphi_t(t)$ on $[\alpha,\beta]$. The last equality means that $F(\varphi(t))$ is an antiderivative of $f(\varphi(t))\cdot\varphi_t(t)$ on $[\alpha,\beta]$. Then the [Newton-Leibniz formula]

$\int_\alpha^\beta f(\varphi(t))\cdot\varphi_t(t)dt = F(\varphi(\beta))-F(\varphi(\alpha)) = F(b)-F(a)$ **[N1].** In the same time for the integral

$\int_a^b f(x)dx$, according to the same [Newton-Leibniz formula], we also have

$\int_a^b f(x)dx = F(b)-F(a)$ **[N2]**. From **[N1]** and **[N2]** follows the equality we need.

Def. For any function $\eta(x)$ on $[a,b]$, the difference $\eta(b)-\eta(a)$ can be denoted as $\eta(x)|_a^b \equiv \eta(b)-\eta(a)$.

Theorem5 [Integration by parts]. $f(x)$ and $g(x)$ are smooth functions on $[a,b]$.

Then the next formula is true: $\int_a^b f(x)\cdot g_x(x)dx = f(x)\cdot g(x)|_a^b - \int_a^b f_x(x)\cdot g(x)dx$.

And the [**Newton-Leibniz formula**] can be rewritten as $\int_a^b f(x)dx = F(x)|_a^b$.

Proof. Let's consider the function $f(x)\cdot g(x)$, this function is differentiable on $[a,b]$ and (according to the standard formula) $(f(x)\cdot g(x))_x = f_x(x)\cdot g(x) + f(x)\cdot g_x(x)$ the right part of the last equality is a continuous function on $[a,b]$, then the same is true for the left part. Then the next integrals are defined and equal $\int_a^b (f(x)\cdot g(x))_x dx = \int_a^b (f_x(x)\cdot g(x) + f(x)\cdot g_x(x))dx \Rightarrow$

$$\Rightarrow f(x)\cdot g(x)|_a^b = \int_a^b (f_x(x)\cdot g(x) + f(x)\cdot g_x(x))dx \Rightarrow$$

$$\Rightarrow f(x)\cdot g(x)|_a^b = \int_a^b f_x(x)\cdot g(x)dx + \int_a^b f(x)\cdot g_x(x)dx \Rightarrow \int_a^b f(x)\cdot g_x(x)dx = f(x)\cdot g(x)|_a^b - \int_a^b f_x(x)\cdot g(x)dx$$

We have built the theory of integral and proved the most important theorems.

We will not consider the simpliest examples like $\int_0^{\pi/2} \sin x dx = -\cos x|_0^{\pi/2} = 1$, because such examples are in abundance in many books and internet resources. There will be several good exercises for Reader's practice.

For Reader's practice:

[1] Show that $0 < \int_{12}^{20} \frac{\cos^2 x}{1+x^8} dx < \frac{1}{10^7}$. **Solution.** For any $x \in [12,20]$ we have

$$0 \le \frac{\cos^2 x}{1+x^8} < \frac{1}{1+12^8} < \frac{1}{12^8} \text{ then } 0 < \int_{12}^{20} \frac{\cos^2 x}{1+x^8} dx < \int_{12}^{20} \frac{1}{12^8} dx = \frac{1}{12^8}(20-12) = \frac{8}{12^8} < \frac{1}{10^7}.$$

[2] $f(x)$ is defined and continuous on $[0,1]$. Find the next limit $\lim_{n\to\infty} \sqrt[n]{f\left(\frac{1}{n}\right)\cdot f\left(\frac{2}{n}\right)\cdot\ldots\cdot f\left(\frac{n}{n}\right)}$.

Answer: $e^{\left(\int_0^1 \ln(f(x))dx\right)}$.

[3] Calculate $\lim_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{2n}\right)$. **Answer**: $\ln 2$.

[4] Calculate the integral $\int_0^{\pi}\frac{\sin x}{1+\cos^2 x}dx$. **Answer**: $\frac{\pi^2}{4}$.

[5] Calculate the integral $\int_{-\pi}^{\pi}\sqrt[3]{\sin x}dx$. **Answer**: 0.

[6] Calculate the integral $\int_0^{\pi/2}\frac{1}{2-\sin x}dx$. **Answer**: $\frac{2\pi\sqrt{3}}{9}$.

[7] Calculate the integral $\int_1^{e}\frac{1}{x(1+\ln^2 x)}dx$. **Answer**: $\frac{\pi}{4}$.

[8] Calculate the integral $\int_0^{1}\arcsin\sqrt{x}dx$. **Answer**: $\frac{\pi}{4}$.

[9] Calculate the integral $\int_0^{\pi/2}\frac{1}{1+\sin x+\cos x}dx$. **Answer**: $\ln 2$.

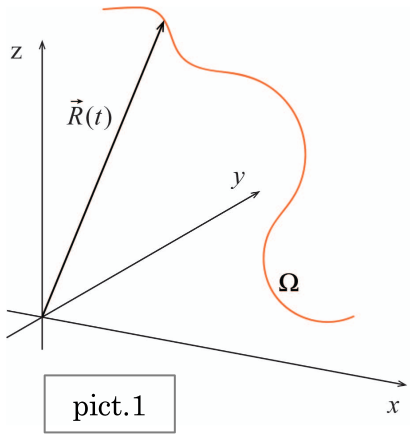
[10] Calculate the integral $\int_1^{2}\frac{e^{1/x^2}}{x^3}dx$. **Answer**: $(e-e^{1/4})/2$.

[11] Calculate the integral $\int_0^{1}\sqrt{4-x^2}dx$. **Answer**: $\frac{\pi}{3}+\frac{\sqrt{3}}{2}$.

Curves

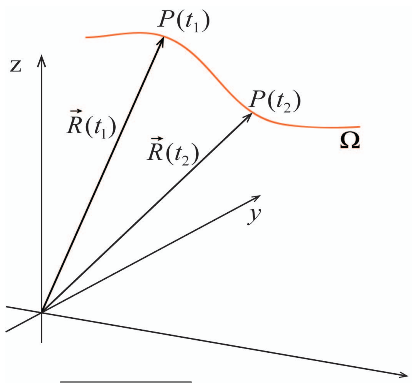
Def1. $x(t), y(t), z(t)$ are some functions on $[a,b]$, then $\vec{R}(t) = (x(t), y(t), z(t)) \| t \in [a,b]$ is called a vector function on $[a,b]$.

For any $t \in [a,b]$ the value $\vec{R}(t)$ is a vector and let's agree to draw it as a radius vector **[pict1]**. The set Ω of points in the space with coordinates $\{(x(t), y(t), z(t)) \| t \in [a,b]\}$ is called a hodograph of $R(t)$ (it is exactly the set of end points of radius vectors $\vec{R}(t) \| t \in [a,b]$). The variable t is called a parameter.
For any $t \in [a,b]$ the point in the space with coordinates $(x(t), y(t), z(t))$ is denoted like $P(t)$.



pict.1

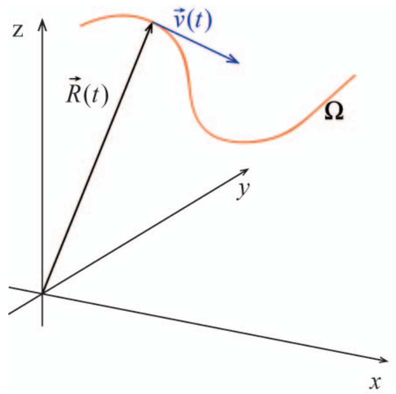
Def2. $\vec{R}(t) = (x(t), y(t), z(t)) \| t \in [a,b]$ is a vector function.
And **[A]** $x(t), y(t), z(t)$ are smooth functions on $[a,b]$ (i.e., these functions are differentiable on $[a,b]$ and their derivatives are continuous functions on $[a,b]$), and
[B] $\forall t_1 \neq t_2 \in [a,b] \Rightarrow P(t_1) \neq P(t_2)$ **[pict2]** (i.e., different values of parameter define different points in the space), and
[C] for any parameter $t \in [a,b]$ the vector $\vec{v}(t) \equiv (x_t(t), y_t(t), z_t(t))$ is a non-zero vector, i.e., $\vec{v}(t) \neq \vec{0}\ \ \forall t \in [a,b]$. Then the hodograph Ω of $\vec{R}(t)$ is called a **smooth curve**, or just a **curve**.



pict.2

A vector $\vec{v}(t) \equiv (x_t(t), y_t(t), z_t(t))$ is usually drawn as a vector, which start point is $P(t)$ **[pict3]** (we will show later that such vector is a "tangent vector" to a curve at a point $P(t)$).

$P(a)$ is called a start point of a curve and $P(b)$ is called an end point of a curve.



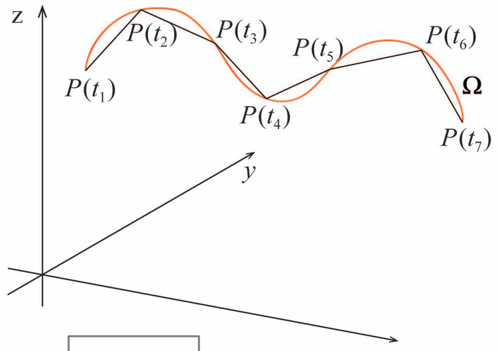
pict.3

From the **def2** follows that for any point P on Ω there exist the unique parameter $t \in [a,b]$ such that $P = P(t)$.
In particular, for any different points $P \neq D$ on Ω, the parameters $t, d \in [a,b]$, where $P = P(t),\ D = D(d)$, are different $t \neq d$.

Let's make any partition of $[a,b]$, so $a = t_1 < t_2 < t_3 < t_4 < ..... < t_n = b$.

Any such partition defines some points $A = P(t_1), P(t_2), P(t_3), ... \ P(t_n) = B$ on the curve Ω.

Let's consider the polygonal chain $P(t_1)P(t_2),\ P(t_2)P(t_3),\ P(t_3)P(t_4),\dots.\ P(t_{n-1})P(t_n)$ **[pict4]**, it's length (by definition) is the sum of lengths of all segments: $L \equiv P(t_1)P(t_2)+P(t_2)P(t_3)+P(t_3)P(t_4)+\ldots+P(t_{n-1})P(t_n)$.



pict.4

Def3. Ω is some curve. If there exist some real number $L(\Omega)$ such that for any (small) positive $\varepsilon>0$ there exist the positive $\delta>0$ such that for any partition $a=t_1<t_2<t_3<t_4<\dots..<t_n=b$ of $[a,b]$, which norm is less than δ (i.e., $\max(t_{k+1}-t_k)<\delta$), we have $|L(\Omega)-L|<\varepsilon$, then $L(\Omega)$ is called a length of a curve Ω.

In the **def3** L is a length of a polygonal chain $P(t_1),P(t_2),P(t_3),\dots.\ P(t_n)$.

Theorem1. If such number $L(\Omega)$ exists, then $L(\Omega)$ is a supremum of $\{L\}$, where $\{L\}$ is a set of lengths of all possible polygonal chains which correspond to all possible partitions: $a=t_1<t_2<t_3<t_4<\dots..<t_n=b$ of $[a,b]$.

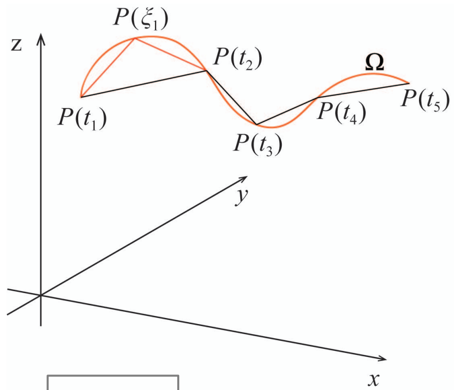
Consequence. The length $L(\Omega)$ is greater than the length L of any polygonal chain (which corresponds to some partition of $[a,b]$), in the same time, the length L of a chain can approach arbitrary close $L(\Omega)$.

Proof. **[Step1]** Let's show that the length L of any polygonal chain is not greater than $L(\Omega)$.

Let's assume that it is not true, then there exist some chain $\tilde{L}$ such that $L(\Omega)<\tilde{L}$, such chain corresponds to some concrete partition $a=t_1<t_2<t_3<t_4<\dots..<t_n=b$ of $[a,b]$.

Let $\tilde{\varepsilon}$ is a distance between numbers $L(\Omega)$ and $\tilde{L}$, i.e., $\tilde{\varepsilon}\equiv\tilde{L}-L(\Omega)>0$.

Let's notice that if we add any one new point ξ_1 to our partition $a=t_1<t_2<t_3<t_4<\dots..<t_n=b$, we will get a new partition, and the length L of a polygonal chain, that corresponds to the new partition, is not less than the length of the chain $\tilde{L}$ that corresponds to the initial partition. Really, let's add one new point ξ_1 between t_1,t_2 **[pict5]**, then we have a new partition $a=t_1<\xi_1<t_2<t_3<t_4<\dots..<t_n=b$, obviously $P(t_1)P(t_2)\le P(t_1)P(\xi_1)+P(\xi_1)P(t_2)$ (triangle inequality). Then $L\ge\tilde{L}$.



pict.5

Let's take now the positive number $\tilde{\varepsilon}/2>0$, there exist δ such that for any partition of $[a,b]$, which norm is less than δ, the length of the chain L which corresponds to such partition, satisfies to $|L(\Omega)-L|<\tilde{\varepsilon}/2$.

Let's refine the initial partition $a=t_1<t_2<t_3<t_4<.....<t_n=b$ by adding some new points ξ_m between t_k, t_{k+1}, in order to get the partition which norm is less than δ. The chain L, which corresponds to such partition, has the property $|L(\Omega)-L|<\tilde{\varepsilon}/2$ and $L\geq\tilde{L}$.

Let's sum up: we have **[1]** $L(\Omega)<\tilde{L}$ and $\tilde{\varepsilon}=\tilde{L}-L(\Omega)$ and in the same time

[2] $|L(\Omega)-L|<\tilde{\varepsilon}/2$ and $\tilde{L}\leq L$.

It's easy to mark the numbers $L(\Omega), L, \tilde{L}$ on some coordinate line (by taking into account **[1]**,**[2]**) and to see that we have a contradiction.

Then the initial assumption $L(\Omega)<\tilde{L}$ was false and the **[step1]** is done.

[Step2] Let's show that for any positive $\varepsilon>0$ the half interval $(L(\Omega)-\varepsilon, L(\Omega)]$ contains at least one element from the set $\{L\}$ (where $\{L\}$ is a set of lengths of all possible polygonal chains).

Let's fix any $\varepsilon>0$, then there exist $\delta>0$ such that **for any** partition, which norm is less than δ, we have $|L(\Omega)-L|<\varepsilon$. So let's fix any such partition, for the chain $\tilde{L}$, which corresponds to it, we have $|L(\Omega)-\tilde{L}|<\varepsilon$ and $\tilde{L}<L(\Omega)$ from here follows that $\tilde{L}\in(L(\Omega)-\varepsilon, L(\Omega)]$.

From the **[Step1]** and **[Step2]** follows that $L(\Omega)$ is a supremum of $\{L\}$.

Consequence1. From the **theorem1** follows that if a curve Ω has a length $L(\Omega)$, then the number $L(\Omega)$ is uniquely defined. Really, if $L(\Omega)$ exists, then it is a supremum of some concrete set $\{L\}$, and any set may have only one supremum.

Theorem2. For any curve Ω the number $L(\Omega)$ exists and $L(\Omega)=\int_a^b\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}\,dt$

Proof. Let's notice that the integral $\int_a^b\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}\,dt$ do exist, because $x(t), y(t), z(t)$ are smooth on $[a,b]$(**def2 [A]**), then $[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2$ is continuous on $[a,b]$ and even a strictly positive on $[a,b]$ (**def2 [C]**), then the composite function $\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}$ is continuous on $[a,b]$ and therefore it is integrable on $[a,b]$.

Let's fix an any partition $a=t_1<t_2<...<t_n=b,$ this partition defines the polygonal chain with

length $L = P(t_1)P(t_2) + P(t_2)P(t_3) + P(t_3)P(t_4) + \ldots + P(t_{n-1})P(t_n) \Leftrightarrow$

$\Leftrightarrow L = \sum_{k=1}^{n-1} \sqrt{(x(t_{k+1}) - x(t_k))^2 + (y(t_{k+1}) - y(t_k))^2 + (z(t_{k+1}) - z(t_k))^2}$. Let's notice that we have here the differences of $x(t), y(t), z(t)$ at the end points of segments $[t_k, t_{k+1}]$, then we can apply the mean value theorem. Let's consider $x(t)$ on $[t_k, t_{k+1}]$, there exist some point $h_k \in (t_k, t_{k+1})$ such that $x(t_{k+1}) - x(t_k) = x_t(h_k)(t_{k+1} - t_k)$ and similarly for the other functions $y(t), z(t)$.

$$L = \sum_{k=1}^{n-1} \sqrt{[x_t(h_k)]^2 \cdot (t_{k+1} - t_k)^2 + [y_t(p_k)]^2 \cdot (t_{k+1} - t_k)^2 + [z_t(u_k)]^2 \cdot (t_{k+1} - t_k)^2} =$$

$= \sum_{k=1}^{n-1} \sqrt{[x_t(h_k)]^2 + [y_t(p_k)]^2 + [z_t(m_k)]^2} \cdot (t_{k+1} - t_k)$ **[S1]** -this sum looks very similar to the integral sum **[S2]** of $\sqrt{[x_t(t)]^2 + [y_t(t)]^2 + [z_t(t)]^2}$ on $[a,b]$.

[S2] $\Sigma = \sum_{k=1}^{n-1} \sqrt{[x_t(\xi_k)]^2 + [y_t(\xi_k)]^2 + [z_t(\xi_k)]^2} \cdot (t_{k+1} - t_k)$ (here $\xi_k \in [t_k, t_{k+1}]$). But anyway, the sum **[S1]** is different, because points h_k, p_k, m_k may be different, and they are determined by the **mean value theorem**, these are some points which belong to the same segment $[t_k, t_{k+1}]$. And in the sum **[S2]** we have only one arbitrary point ξ_k on $[t_k, t_{k+1}]$, so **[S1]** and **[S2]** are really different sums.

Let's fix an arbitrary positive $\bar{\varepsilon} > 0$.

[Step1] We will use inequality **[E]** $\left|\sqrt{a} - \sqrt{b}\right| \le \sqrt{|a-b|}$ $\forall a \ge 0, b \ge 0$.

As $x_t(t), y_t(t), z_t(t)$ are continuous on $[a,b]$, then $[x_t(t)]^2, [y_t(t)]^2, [z_t(t)]^2$ are also continuous on $[a,b]$, then these functions are uniformly continuous on $[a,b]$. Then the for the positive number $\frac{\bar{\varepsilon}^2}{12(b-a)^2}$ there exist $\bar{\delta} > 0$ such that for any $t_1, t_2 \in [a,b]$, as soon as $|t_1 - t_2| < \bar{\delta}$, then, right away:

$$\left|[x_t(t_1)]^2 - [x_t(t_2)]^2\right| < \frac{\bar{\varepsilon}^2}{12(b-a)^2}, \quad \left|[y_t(t_1)]^2 - [y_t(t_2)]^2\right| < \frac{\bar{\varepsilon}^2}{12(b-a)^2}, \quad \left|[z_t(t_1)]^2 - [z_t(t_2)]^2\right| < \frac{\bar{\varepsilon}^2}{12(b-a)^2}.$$

Let's fix an arbitrary partition of $[a,b]$ which norm is less than $\bar{\delta}$, we need to estimate how close is the sum **[S1]** to the sum **[S2]**.

$$|L - \Sigma| = \left|\sum_{k=1}^{n-1} \left(\sqrt{[x_t(h_k)]^2 + [y_t(p_k)]^2 + [z_t(m_k)]^2} - \sqrt{[x_t(\xi_k)]^2 + [y_t(\xi_k)]^2 + [z_t(\xi_k)]^2}\right) \cdot (t_{k+1} - t_k)\right| \le [\textit{inequality } [E]]$$

$$\le \left|\sum_{k=1}^{n-1}\left(\sqrt{\left|\ [x_t(h_k)]^2-[x_t(\xi_k)]^2+[y_t(p_k)]^2-[y_t(\xi_k)]^2+[z_t(m_k)]^2-[z_t(m_k)]^2\ \right|}\right)\cdot(t_{k+1}-t_k)\right|\le$$

$$\le \left|\sum_{k=1}^{n-1}\left(\sqrt{\left|[x_t(h_k)]^2-[x_t(\xi_k)]^2\right|+\left|[y_t(p_k)]^2-[y_t(\xi_k)]^2\right|+\left|[z_t(m_k)]^2-[z_t(m_k)]^2\right|}\right)\cdot(t_{k+1}-t_k)\right|<$$

$$<\left|\sum_{k=1}^{n-1}\left(\sqrt{\frac{\bar\varepsilon^2}{12(b-a)^2}+\frac{\bar\varepsilon^2}{12(b-a)^2}+\frac{\bar\varepsilon^2}{12(b-a)^2}}\right)\cdot(t_{k+1}-t_k)\right|=\sum_{k=1}^{n-1}\frac{\bar\varepsilon}{2(b-a)}\cdot(t_{k+1}-t_k)=\frac{\bar\varepsilon}{2(b-a)}\cdot(b-a)=\bar\varepsilon/2$$

[Step2] From the main definition of integral $\mathrm{A}\equiv\int_a^b\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}\,dt$ follows that:

for $\bar\varepsilon/2>0$ there exist some positive $\tilde\delta\le\bar\delta$ such that for any partition of $[a,b]$, which norm is less than $\tilde\delta$, we have $|\mathrm{A}-\Sigma|<\bar\varepsilon/2$ (here Σ is an integral sum **[S2]**). And from the **[step1]**, for any such partition there also must be $|L-\Sigma|<\bar\varepsilon/2$.

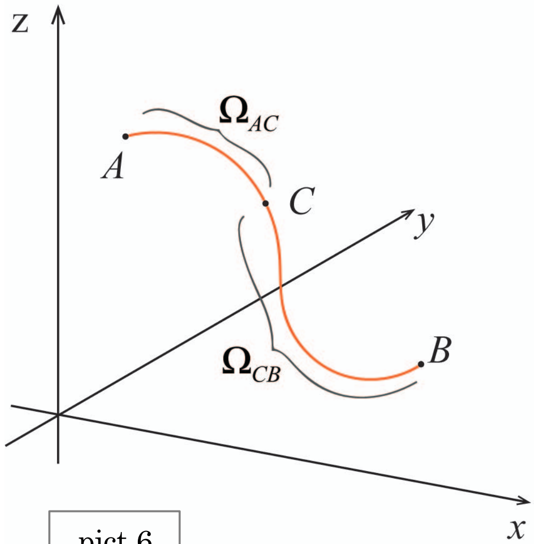
Let's sum up: from $|L-\Sigma|<\bar\varepsilon/2$ and $|\mathrm{A}-\Sigma|<\bar\varepsilon/2$ follows that $|L-\mathrm{A}|<\bar\varepsilon$.

We had started from an arbitrary positive $\bar\varepsilon$, and we showed that there exist some $\tilde\delta>0$ such that for any partition, which norm is less than $\tilde\delta$, we have $|L-\mathrm{A}|<\varepsilon$.

Then, according to the def3, A is a length of the curve Ω. Everything is proved.

Consequence2. Any point C on Ω divides Ω in two curves Ω_{AC} and Ω_{CB} **[pict6]**.

From the simpliest properties of integral and theorem2 we have: $L(\Omega)=L(\Omega_{AC})+L(\Omega_{CB})$.



pict.6

Let's take any $c<d\in[a,b]$. The length of the curve with end points $P(c)$ and $P(d)$ is equal to

$\int_c^d\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}\,dt$ **[F]**. The function

$\varphi(t)\equiv\sqrt{[x_t(t)]^2+[y_t(t)]^2+[z_t(t)]^2}$ is continuous on $[a,b]$

and therefore it is continuous on $[c,d]$, then it reaches it's maximum $\mathrm{M}_{[c,d]}$ and it's minimum $m_{[c,d]}$ at some points of the segment $[c,d]$. From the def2 **[C]** follows that $\varphi(t)$ is strictly positive on $[c,d]$. As there exist some points $u,z\in[c,d]$ such that $\varphi(u)=\mathrm{M}_{[c,d]}$ and $\varphi(z)=m_{[c,d]}$, then $\mathrm{M}_{[c,d]}>0$ and $m_{[c,d]}>0$. So we have

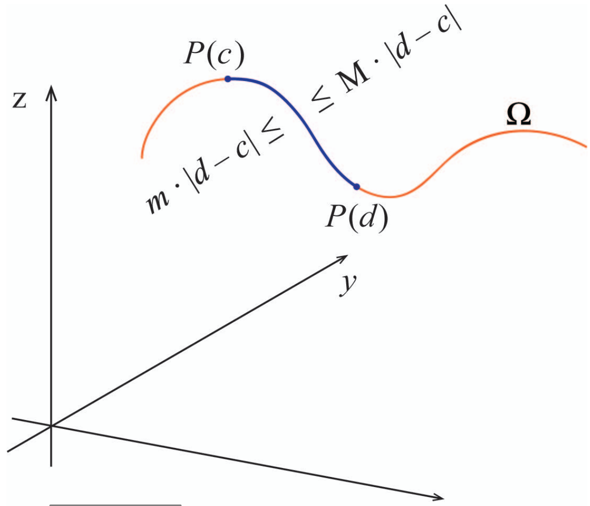
$0<m_{[c,d]}\le\varphi(t)\le\mathrm{M}_{[c,d]}\,||\,\forall t\in[c,d]$. And obviously $0<m_{[a,b]}\le m_{[c,d]}$ and $\mathrm{M}_{[c,d]}\le\mathrm{M}_{[a,b]}$.

Then $0 < m_{[a,b]} \le \varphi(t) \le \mathrm{M}_{[a,b]} \,||\, \forall t \in [c,d]$. Let's designate $m_{[a,b]} \equiv m$ and $\mathrm{M}_{[a,b]} \equiv \mathrm{M}$. Then we have $0 < m \le \varphi(t) \le \mathrm{M} \,||\, \forall t \in [c,d]$. Let's estimate the integral **[F]** (the length of the curve between $P(c)$ and $P(d)$): $0 < m\cdot(d-c) \le \int_c^d \sqrt{[x_t(t)]^2 + [y_t(t)]^2 + [z_t(t)]^2}\,dt \le \mathrm{M}\cdot(d-c)$ **[V]**.

From **[V]** follows that: if the values of parameters $c < d$ are very close, then the length of the curve between $P(c)$ and $P(d)$ is a very small number. And conversely, if the length of the curve between $P(c)$ and $P(d)$ is a small number, then c must be close to d. We got **[V]** from the assumption $c < d \in [a,b]$.

If $d < c \in [a,b]$ we can get the similar formula.

And we can unite these results in one: for any $c,d \in [a,b] \,||\, c \ne d$, the length of the curve between $P(c)$ and $P(d)$ is $\left|\int_c^d \sqrt{[x_t(t)]^2 + [y_t(t)]^2 + [z_t(t)]^2}\,dt\right|$ and



pict.7

the next estimation is true $0 < m\cdot|d-c| \le \left|\int_c^d \sqrt{[x_t(t)]^2 + [y_t(t)]^2 + [z_t(t)]^2}\,dt\right| \le \mathrm{M}\cdot|d-c|$ **[U] [pict7]**

(it's very easy to check that this estimation is true in any case: $c < d,\ d < c$, the case $c = d$ is not allowed, we do not need it. And our reasoning was done for different values of parameter c,d).

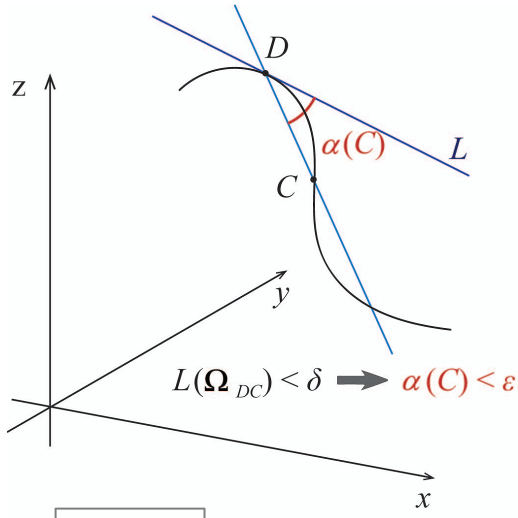
And obviously, for any $\tau \in (a,b]$ the length of the curve between $P(a),\ P(\tau)$ is $L(\tau) \equiv \int_a^\tau \sqrt{[x_t(t)]^2 + [y_t(t)]^2 + [z_t(t)]^2}\,dt$. As we have a strictly positive function inside the integral, then $L(\tau)$ is a strictly increasing function on $[a,b]$. (it means that when τ moves from a to b, the length of the curve between $P(\tau)$ and $P(a)$ grows, which makes sense).

Exercise1. Show that $L(\tau)$ is continuous on $[a,b]$ (use the estimation **[U]**).

Conclusion. The length function $L(\tau)$ is strictly increasing and continuous on $[a,b]$.

From now on, the points on a curve, which correspond to parameters c,d, we denote as C,D (i.e., $C \equiv P(c),\ D \equiv P(d)$).

Def4. Ω is a curve and D is a fixed point on Ω.
The line L, which passes through the point D,
is called a **tangent line** to Ω at the point D if:
for any (small) positive $\varepsilon > 0$ there exist $\delta > 0$ such
that for any point C on the curve, where
$0 < L(\Omega_{DC}) < \delta$, the angle $\alpha(C) \equiv \angle(L, DC) < \varepsilon$, i.e.,
the angle between L and DC (in radians)
is less than ε **[pict8]**.



pict.8

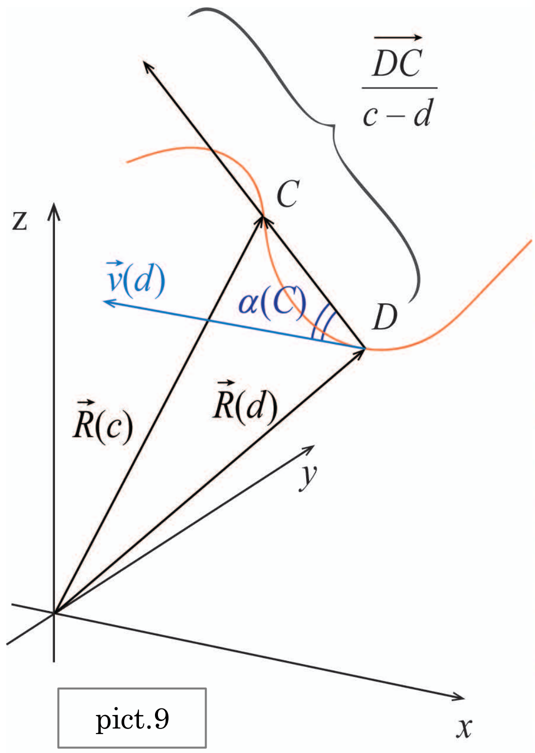
Exerise2. From the def4 follows that if there exist a tangent line L at some point D, then it is unique (there can't be two different tangent lines at the same point D).

Def5. L is a line, any non-zero vector $\vec{v}$, which lies on L, is called a direction vector of $P(d)$.

Theorem3. For any point $P(d)$ on Ω, the tangent line L at $P(d)$ exists.
And $\vec{v}(d) \equiv (x_t(d), y_t(d), z_t(d))$ is a direction vector of L.

Proof. As there can be only one tangent line at any point, it's enough to show that the line L, which passes through D and which direction vector is $\vec{v}(d) \equiv (x_t(d), y_t(d), z_t(d))$, is exactly the line we need (we will show that this line L satisfies to the def4 and therefore L is a tangent line).
Let's fix an arbitrary positive $\varepsilon > 0$ and any point D on the curve. The line DC goes through the points $D, C \in \Omega$, so we can call it "the secant DC". And $\overrightarrow{DC}$ is a direction vector of the secant DC.

Obviously $\vec{R}(d) + \overrightarrow{DC} = \vec{R}(c) \Rightarrow \overrightarrow{DC} = \vec{R}(c) - \vec{R}(d)$. We have $\vec{R}(c) = (x(c), y(c), z(c))$ and $\vec{R}(d) = (x(d), y(d), z(d))$, then $\overrightarrow{DC} = (x(c) - x(d), y(c) - y(d), z(c) - z(d))$ as C and D are different points on the curve, then c and d must be different numbers, therefore the number $\dfrac{1}{c-d}$ is defined (it can be a positive or a negative number, it doesn't matter).
Let's consider the vector which is proportional to $\overrightarrow{DC}$,
the vector $\dfrac{\overrightarrow{DC}}{c-d} = \left(\dfrac{x(c)-x(d)}{c-d}, \dfrac{y(c)-y(d)}{c-d}, \dfrac{z(c)-z(d)}{c-d} \right)$
it is still a direction vector of the secant DC.
The angle $\alpha(C)$ is the angle between vectors **[pict9]**
$\vec{v}(d) \equiv (x_t(d), y_t(d), z_t(d))$ and



pict.9

$$\frac{\overrightarrow{DC}}{c-d}=\left(\frac{x(c)-x(d)}{c-d},\frac{y(c)-y(d)}{c-d},\frac{z(c)-z(d)}{c-d}\right).$$

Notice that $\vec{v}(d)$ is a constant vector, all it's components are fixed numbers. And $\frac{\overrightarrow{DC}}{c-d}$ is not a constant vector, it's position is completely defined by the parameter c (and d is a fixed parameter, because D is fixed).

Let's write the dot product of vectors $\vec{v}(d)$ and $\frac{\overrightarrow{DC}}{c-d}$ in two different ways:

$$|\vec{v}(d)|\cdot\left|\frac{\overrightarrow{DC}}{c-d}\right|\cdot\cos(\alpha(C))=x_t(d)\cdot\frac{x(c)-x(d)}{c-d}+y_t(d)\cdot\frac{y(c)-y(d)}{c-d}+z_t(d)\cdot\frac{z(c)-z(d)}{c-d}\Rightarrow$$

$$\Rightarrow\cos(\alpha(C))=\frac{x_t(d)\cdot\frac{x(c)-x(d)}{c-d}+y_t(d)\cdot\frac{y(c)-y(d)}{c-d}+z_t(d)\cdot\frac{z(c)-z(d)}{c-d}}{\sqrt{[x_t(d)]^2+[y_t(d)]^2+[z_t(d)]^2}\cdot\sqrt{\left[\frac{x(c)-x(d)}{c-d}\right]^2+\left[\frac{y(c)-y(d)}{c-d}\right]^2+\left[\frac{z(c)-z(d)}{c-d}\right]^2}}\quad\textbf{[T]}$$

The expression on the right side of **[T]** is a function of a variable c that is defined for any $c\in[a,d)\cup(d,b]$. It's easy to understand (from the simpliest properties of limits) that the right part of **[T]** has a limit 1 when $c\to d$, this expression is a function of a variable c, let's denote it $\theta(c)$. As $\theta(c)$ is always a cosine of some angle, we always have $-1\le\theta(c)\le 1$.

Let's sum up: we have **[T]** $\cos(\alpha(C))=\theta(c)\,||\,-1\le\theta(c)\le 1\,||\,c\in[a,d)\cup(d,b]$, then we can take the $\arccos()$ of both sides of the first equality of **[T]**,

$\arccos(\cos(\alpha(C)))=\arccos(\theta(c))\Leftrightarrow\alpha(C)=\arccos(\theta(c))$. Then $\alpha(C)$ is a composite function $\arccos(\theta(c))$. As $\lim_{c\to d}\theta(c)=1$ and $\arccos y$ is left continuous at $y=1$, then (theorem about a limit of a composite function) the limit $\lim_{c\to d}\arccos(\theta(c))$ exists and equal to $\arccos 1=0$, then $\lim_{c\to d}\alpha(C)=0$.

Let's fix an arbitrary positive $\bar{\varepsilon}>0$, there exist the positive $\bar{\delta}$ such that as soon as $0<|c-d|<\bar{\delta}$ we have $|\alpha(C)-0|<\bar{\varepsilon}\Leftrightarrow\alpha(C)<\bar{\varepsilon}$ (because $\alpha(C)$ is always a positive angle).

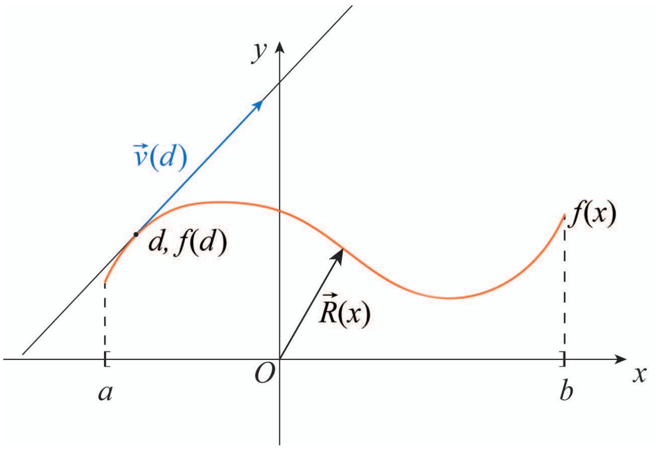
We have the estimation **[U]** $0<m\cdot|d-c|\le L(\Omega_{CD})\le \mathrm{M}\cdot|d-c|$. Let's take the positive number $\bar{\delta}\cdot m$, then for any point C on the curve, for which $L(\Omega_{CD})<\bar{\delta}\cdot m$, then we have $m\cdot|d-c|\le\bar{\delta}\cdot m\Rightarrow|d-c|<\bar{\delta}$ and therefore $\alpha(C)<\bar{\varepsilon}$.

So, for any positive $\bar{\varepsilon} > 0$, we can always find the positive $\delta \equiv \bar{\delta} \cdot m$ such that if $L(\Omega_{CD}) < \bar{\delta} \cdot m$, then $\alpha(C) < \varepsilon$.

We showed that the line L, which direction vector is $\vec{v}(d) \equiv (x_t(d), y_t(d), z_t(d))$, satisfies to the **def4**, then L is a tangent line.

A graph of any smooth function $f(x)$ on some segment $[a,b]$ can be considered as a **2**-dimentional curve Ω, i.e., **[pict10]** the graph of $f(x)$ on $[a,b]$ is the same as a hodograph of the vector function $\vec{R}(x) \equiv (x, f(x)) \| x \in [a,b]$. Notice that the graph of $f(x)$ is really a curve, because it satisfies to the **def2**.

Let's fix any point D on the graph of $f(x)$, it has some coordinates $(d, f(d))$.



pict.10

According to the **theorem3**, there exist a tangent line L at this point, and it's direction vector is $\vec{v}(d) = (1, f_x(d))$. In general, let L is a line on the plane which direction vector is $(a,b) \| a \neq 0$.

Then for the angle α between L and the positive direction of Ox we have $tg\alpha = b/a$ (the number $tg\alpha$ is called a slope of a line L). In our case, the direction vector is $\vec{v}(d) = (1, f_x(d))$, then we have the slope $tg\alpha = f_x(d)$. Then the derivative of a smooth function $f(x) \| x \in [a,b]$ at any point $d \in [a,b]$ is a slope of a tangent line (to the graph of f) at the point $(d, f(d))$.

Exercise3. Does there exist any curve which passes through every point of some unit cube?

Hint. Show that any curve has zero volume, i.e., for any $\varepsilon > 0$ it can be covered by several rectangular parallelepipeds, which total volume is less than ε.

[1] Find the length of the curve:

$x = (t^2 - 2)\sin t + 2t\cos t,\ \ y = (t^2 - 2)\cos t - 2t\sin t \| t \in [0, \pi]$. **Answer:** $\pi^3/3$

[2] Find the length of the curve:

$x = \cos^3 t,\ \ y = \sin^3 t,\ \ z = \cos 2t \| t \in [0, 2\pi]$. **Answer:** 10

[3] Find the length of the curve: $y = \ln \sin x \| x \in [\pi/3,\ \pi/2]$. **Answer:** $\ln 3/2$

[4] Find the length of the curve: $x = \cos t + t\sin t,\ \ y = \sin t - t\cos t \| x \in [0,\ \pi/3]$. **Answer:** $\pi^2/18$

Literature

1. Course of lectures of mathematical analysis.Part 1. V.I. Kolyada, A.A. Korenovskii. Odessa, 2009.
2. Differential geometry, E.G. Poznyak. MGU, 1990.
3. Course of differential and integral calculus,Volume 1. G.M. Fihtengolz. FizMatLit, 2001.
4. Linear algebra, E.S. Kochetkov, A.V. Osokin. Moscow, 2012.
5. Linear Algebra. A.I. Kozko. MGU, 2018
6. Lectures on algebra, D.K. Fadeev, Moscow, 1984.
7. Mathematical Analysis. A.I.Kozko. MGU 2017
8. Course of mathematical analysis. A.M. Ter-Krikorov, M.I. Shabunin. FizMatLit, 2001.
9. Numbers and polynomials. I.V Proskuryakov. Moscow, 1965.
10. Mathematical analysis, Part I. M.V. Falaleev. Irkutsk, 2013.
11. Collection of problems in mathematical analysis. Parts 1,2. L.D.Kudriavcev, A.D.Kutasov, V.I.Chehlov, M.I.Shabunin. Fizmatlit, 2003.
12. Mathematics for Jee Main & Advanced. Book2, Book4. Wiley, Plancess. Mumbai 2015.
13. Encyclopedia of elementary mathematics, Volume 4, P. Aleksandrov, A. Markushevich, A. Hinchin. Leningrad, 1953.
14. Multiple and curvilinear integrals, G.M. Fihtengolz.
15. Encyclopedia of elementary mathematics, Volume 5. P. Aleksandrov, A. Markushevich, A. Hinchin. Leningrad, 1966.
16. The video "Complex Numbers as Matrices". Youtube, channel Mathoma.
17. Technique of calculating limits, G. S. Shukova. RHTU, 2005.
18. Theory of functions of a complex variable in problems and exercises, L.N. Posicelskaya, Moscow, 2007.